FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Atiq Mahmud

PREDICTION OF THE HRV SIGNAL DURING TREADMILL RUNNING

Master's Thesis
Degree Programme in Biomedical Engineering: Signal & Image Processing
December 2019

Mahmud A. (2019) Prediction of the HRV signal during treadmill running. University of Oulu, Degree Programme in Biomedical Engineering: Signal & Image Processing. Master's Thesis, 54 p.

ABSTRACT

Our heart rate is varying every time, and the autonomic nervous system maintains this complex control mechanism. Analysis of heart rate variability (HRV) is a useful tool for autonomic nervous system assessment. It can be a useful marker for different cardiac arrhythmias and heart diseases, and its' clinical relevance is increasing day by day. HRV analysis has an important impact on exercise physiology since it can be a useful marker for stress and recovery. HRV during exercise differs a lot from the normal condition as body movement, exercise intensity, and other factors modulate the HRV. Few recent studies show the effect of running cadence and pedaling frequency on the HRV during treadmill exercise and cycling exercise, respectively. Our research is based on incremental treadmill exercise, and we tried to figure out which part of HRV can be explained by running cadence. We tried to create a polynomial model for HRV, which can predict the future HRV by training the model with appropriate training data and later validate the model with the HRV signal from different running intervals. We observed a significant reduction in the model performance with the increment of running speed. The reduction in model performances validates that the HRV signal is affected most when the running intensity is maximum. We tried to correlate our model residuals with the actual acceleration signal, but due to some complexity, we couldn't achieve what we have hypothesized.

Keywords: Electrocardiogram, Heart rate variability, Autonomic nervous system assessment, Physical exercise, Polynomial model, Autoregressive model, Autoregressive model with exogenous input.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

ANS	Autonomic nervous system
AR	Autoregressive
ARX	Autoregressive model with exogenous input
AV	Atrioventricular
CLC	Cardio-locomotor coupling
CNS	Central nervous system
ECG	Electrocardiogram
HF	High frequency
HR	Heart rate
HRV	Heart rate variability
LA	Left atrium
LF	Low frequency
LV	Left ventricle
RA	Right atrium
RSA	Respiratory sinus arrhythmia
RV	Right ventricles
SA	Sinoatrial
mV	Millivolt
Hz	Hertz
kHz	Kilohertz
$y(t)$	System output
$u(t)$	System input
$v(t)$	Output disturbance or noise
$g(k)$	Discrete impulse response of the system
$h(k)$	Discrete impulse response of the noise system
$e(t)$	White noise error
$A(q)$	Denominator polynomial related to y
$B(q)$	Numerator polynomial related to u
$G(q)$	Rational transfer function related to u
$H(q)$	Rational transfer function related to e
q^{-1}	Backward-shift operator
$\sum$	Sum
ϕ	Regressor vector
$\varepsilon(t)$	Estimated prediction error
θ	Parameter vector
$J(\theta)$	Scalar objective function
n_a	Order of polynomial A(q)
n_b	Order of polynomial B(q)
n_c	Order of polynomial C(q)

1. INTRODUCTION

The autonomic nervous system (ANS) plays a significant role in maintaining the homeostasis of the body by controlling the heart rate regulation and many other vital functions. A healthy heart contains irregularity in the heartbeats. The heart rate is varying all the time, depending on the required cardiac output of the situation, i.e., running, resting, etc. Heart rate variability (HRV) explains the inter-beat differences in the heart rate. In the past few decades, analysis of HRV has gained massive popularity as it's a popular non-invasive method for ANS assessment.

The sympathetic and parasympathetic nervous system controls the heart rate. During exercise, the general assumption is that the sympathetic nervous system controls the heart rate and provide necessary increment on the heart rate based on the exercise intensity. This complex control mechanism is so fast that the heart rate can be doubled within 3 to 5 *seconds*. The sinus node in the heart is the center of this control system as it maintains the heart rate according to the information from the ANS.

In normal conditions, one of the major sources of HRV is the respiratory sinus arrhythmia (RSA), which is the synchronization of the HRV and respiration. During running and other physical exercises, the effect of respiratory modulation is gradually diminishing, and other sources like cadence and its' harmonic start to dominate the HRV. Various studies can be found based on the effect of different physical exercises on the HRV signal. Iman et al. [1] showed the impact of body movement on the HRV signal during treadmill running. This study is mainly focused on the high intensity running as HRV can be most affected by physical movement during intensive exercise. The relation between HRV and body movement is known as cardio-locomotor coupling (CLC), and the authors used time-frequency analysis of the HRV signal to find out the CLC components in the spectrum. They claimed that over 50% of their study population showed the influence of the running cadence on HRV. Blain G. et al. [2] used cycling instead of treadmill running to find the effect of pedaling frequency on HRV. They applied the same time-frequency analysis of the HRV signal and observed the continuous effect of pedaling frequency on HRV. They also indicated that in low-workload due to the spectral aliasing, the pedaling frequency components can affect the low-frequency component of the HRV. David et al. [3] studied on the methodological framework of the HRV analysis during physical exercise. The framework is based on the time-frequency representation of HRV as they claimed that the traditional spectral analysis of HRV failed to assess the non-stationarity of HRV during exercise.

In another study, George M. et al. [4] focused on different prediction and approximation methods for HRV analysis, and how this kind of HRV analysis can be used in classification for clinical purposes. Though this research article is based on clinical uses, it provides some general ideas about different predictions and approximation methods for HRV analysis. The performed prediction methods for HRV analysis are, linear prediction, local exponential prediction, prediction with the method of delays, autoregressive (AR) prediction, and neural network prediction. Similarly, the performed approximation methods for HRV analysis are, local linear

approximation, least-squares approximation, neural network approximation, and wavelet approximation.

The authors were trying to build a line between 'too closely' and 'close enough' prediction or approximation so that the errors provided by the 'close enough' prediction or approximation can be used later in classification purposes to differentiate between normal and patient data. The classification process was not performed in this paper, as the authors only provided a hypothesis regarding that. They used three different pairs of datasets for comparison. Two pairs contain data from healthy persons and data from patients who suffered from heart failure and coronary artery disease. The final pair contains only data from the young people and data from the elderly persons for comparison. As a result, they showed that the error occurred by a prediction or approximation method can be useful for classifying a pair of dataset into healthy and patient (for the first two pairs) or young and elder (for the third pair).

M. P. Tarvainen et al. [5] studied the time-varying autoregressive (AR) modeling with a Kalman smoother algorithm to provide a spectrum estimation of HRV. The control system of HRV is really complicated, as the characteristics of HRV are time-varying. The authors first modeled this nonstationary HRV signal with a time-varying AR model, and then by using a Kalman smoother algorithm, they calculated the model parameters. The Kalman smoother consists of two parts, a Kalman filter and a smoother. Finally, the time-varying estimated spectrum from the AR parameter was decomposed into separate components to examine the high-frequency (HF) and low-frequency (LF) components separately. For this study, the orthostatic test (subject lay supine for five minutes and then stood up for the next five minutes) was selected by the authors. In the results, they showed that how changing in body position effect the dynamics of HRV as HF components were dominating while the subject was laying, and as soon as the subject stood up, the LF components became the dominating one.

Steven W. et al. [6] proposed a novel non-parametric Hammerstein model based predictive control for the heart rate regulation during treadmill exercise. They started their hypothesis by stated that the actual model structure of the human cardiovascular system is too hard to determine, and hence, they used a non-parametric modeling approach. They claimed that the designed system can provide optimal heart rate based on the treadmill running, and it can predict the safer speed and acceleration for treadmill exercise. Stefano et al. [7] studied on the non-linear control techniques for the heart rate regulation during long-duration treadmill exercise.

The traditional time-frequency analysis of HRV can be a complicated non-intuitive task. That's why, in this thesis, we are trying to find an alternate option for analyzing the HRV signal. We are focusing on incremental running exercise and trying to observe how HRV can be affected by running cadence and other body movements. Our main aim is to build a simple polynomial model which can be used as an HRV predictor. We are going to introduce the system identification procedure based on the hypothesis that we can predict the future HRV signal based on other input signals, i.e., acceleration and the previous values of the HRV signal. We are assuming that the modeling approach will need some adaptation to cope with the changes in HRV due to physical movement. We are following the generic system identification procedures and introducing modifications according to the necessity. We are also assuming that if we can build a good enough model for HRV prediction, we can use the residual signal from the model to determine the effect of the running cadence on HRV. Autoregressive

(AR) model and autoregressive model with exogenous input (ARX) are used for the prediction of the HRV signal.

In this study, we are trying to build a simple predictor model for HRV based on the used dataset. We are focusing on the model performances and observing the effect of running intensity on model performances. We are starting our modeling approach with a linear polynomial model, and gradually we will try to improve the model performance by altering different parameters.

This thesis is organized in such a way that chapter 2 summarizes the physiological background related to the HRV. The next section explains the methods for generating the HRV signal from raw ECG and the modeling approaches. The experiment results are shown in the following chapter with a modification of the modeling approach. Chapter 5 shows the overall result for the whole dataset and how far we have achieved our hypothesis, and the thesis is concluded in the next chapter.

2. THEORETICAL BACKGROUND

2.1 Autonomic Nervous System

The autonomic nervous system (ANS) is a portion of the peripheral nervous system which works independently without any conscious or voluntary control. John Newport Langley, a British physiologist, described the functions of ANS in 1921. [8 p.7] Maintenance of the homeostasis is one of the major roles of ANS, and to achieve this, ANS performs different physiological changes in internal and external conditions according to fulfill the demand of the situation. [9] For example, a few minutes of running results in a much higher heart rate with a higher respiration rate while a relaxing state can provide the opposite result. Proper knowledge about ANS is important since the activities of almost every tissue and body organ are influenced by this nervous system. Regulation of heart rate, blood pressure, control of respiration, sweating, etc. are a few of the many outcomes of ANS activity. [10]

The autonomic nervous system has either a full or partial effect on all pathological and physiological conditions of the human body. Various endocrine and exocrine glands, cardiac muscle, and smooth muscle are innervated by the ANS. [9] The intensity and speed of executing any change in visceral function is an amazing characteristic of the ANS. For example, it only needs 3 to 5 *seconds* for the heart rate and 10 to 15 *seconds* for the arterial pressure to be doubled from the normal. [11 p.748] There are various factors that can result in a change in ANS activity, like sleeping, physical activity, aging, etc. The ANS plays a vital role in maintaining normal pathological settings. The autonomic imbalance has been used as a marker of abnormal cases in the pathophysiology like myocardial infarction and congestive heart failure. [12] Regular monitoring of ANS activity can be helpful in the early detection of cardiovascular diseases and developing personalized treatment plans.

The ANS activities are activated and controlled by the centers located at the brain stem, spinal cord, hypothalamus, and a portion of the cerebral cortex; altogether, they are known as homeostatic control centers. The hypothalamus and brain stem monitor and control different important factors of the body, such as heart rate, body temperature, blood pressure, hunger, thirst, and many more. [9] Often, the ANS activities are maintained by the visceral reflexes, which are the way to control visceral activities automatically. During a visceral reflex, the subconscious sensory signal originated from the visceral organ, then entered to the homeostatic control centers and finally, returned to the organ directly to provide the necessary control mechanism subconsciously. [11 p.748]

The baroreceptor reflex is a good example of the autonomic reflex performed by the ANS. Baroreceptors are a sensory receptor located in the arteries to monitor blood pressure. The purpose of baroreceptors is to transmit the sensory impulses to the vasomotor center in the brain stem. The amount of sensory impulse depends on blood pressure. If the blood pressure is low, the amount of sensory impulse to the brain stem is low, and this change results in an adjusted ANS activity in heart and blood vessels

to increase the heart rate and vascular resistance. Finally, the blood pressure is back to the normal level. The opposite phenomenon happens if blood pressure is high. [9]

The transmission of the efferent autonomic signals is managed by two anatomically distinct divisions, the sympathetic nervous system and the parasympathetic nervous system. Both divisions are simultaneously active, i.e., they provide a degree of nervous input to any given tissue all the time. The increased or decreased tissue activity is based on the change in the nervous input. This tonic activity is an important characteristic of ANS, and it helps ANS to regulate tissue function more precisely and accurately. Sympathetic tone and parasympathetic tone describes the basal rate of activity of the two divisions of ANS. [10]

Neurons are the fundamental element of the brain and the nervous system. The pathway of the autonomic nervous system consists of two motor neurons, preganglionic neuron and postganglionic neuron. These two neurons are separated from each other by the autonomic ganglia. Motor neurons transmit impulses away from the CNS, and in the case of ANS, they transmit impulses to the effector organs. The postganglionic neuron synapses the target tissue while the preganglionic neuron synapses the postganglionic neuron first. Preganglionic neurons are originated from the brain and spinal cord. Autonomic ganglia are situated in the head, neck, and abdomen and parallel to the spinal cord. The sympathetic and parasympathetic division of ANS is distinguishable by the location of autonomic ganglia and the source of preganglionic neurons. [8 p.222-223]

The neurons of ANS mainly release two neurotransmitters, acetylcholine, and norepinephrine. Neurotransmitters are the chemical messenger of the body. The nerve fibers are called cholinergic if they secrete acetylcholine, and adrenergic if they secrete norepinephrine. The preganglionic neurons in both sympathetic and parasympathetic nervous systems and almost all postganglionic neurons from the parasympathetic nervous system are cholinergic, while most of the postganglionic neurons from the sympathetic nervous system are adrenergic. So, the nerve ends of parasympathetic and sympathetic nerve fiber release acetylcholine and norepinephrine most of the time. For this reason, acetylcholine and norepinephrine are called parasympathetic transmitters and sympathetic transmitters, respectively, and the sympathetic or parasympathetic effect on an organ depends on them. The neurotransmitters activate their corresponding receptors in the ganglia and the target tissue. Muscarinic and nicotinic are the cholinergic receptors, while alpha receptors and beta receptors are two major types of adrenergic receptors. [11 p.750-751]

The sympathetic nervous system provides a ‘fight-or-flight’ response. The origin of sympathetic preganglionic neurons in the thoracic and lumbar region of the spinal cord. Then it passes through the sympathetic chain and reaches the effector tissue. The sympathetic ganglion chain runs parallel to the spinal cord, where the preganglionic neuron synapses the postganglionic neuron. One preganglionic neuron may synapse many postganglionic neurons since there is a 1:20 ratio of the preganglionic and postganglionic neurons. [10] The activity from the sympathetic nervous system often involves the whole system rather than a discrete organ. Figure 1 shows two divisions of ANS with the affecting organs and tissues. [9]

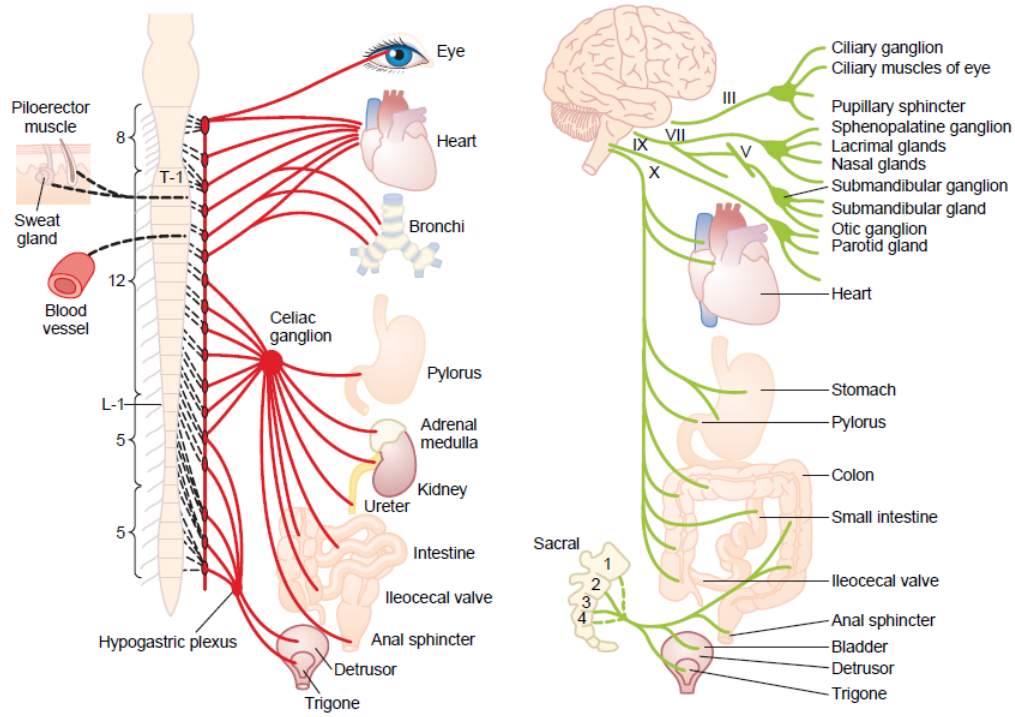

Figure 1. The sympathetic nervous system (left) and the parasympathetic nervous system (right).

The parasympathetic nervous system provides ‘rest-and-digest’ functions since it’s predominating while a person is resting. Brainstem and sacral region of the spinal cord are the originating points for the parasympathetic preganglionic neurons. Since the parasympathetic ganglia are located in the innervated tissue, the parasympathetic preganglionic neurons are much longer, and the number of postganglionic neurons is much lesser than their sympathetic counterparts. Vagus nerve, also known as Cranial Nerve X, is carrying almost 75% of the parasympathetic nerve fibers. [11 p.750] Since both divisions of ANS are active simultaneously, an organ can innervate by both divisions at a time, and the result can be antagonistic, complementary, or cooperative. [8 p.235]

Some major effects of the sympathetic and parasympathetic nervous system to different organ and tissue is shown in Table 1. [11 p.754, 8 p.227] The effect of the sympathetic or parasympathetic nervous system differs from organ to organ, and at least we can say that they show the opposite effect in the same organ. Since our focus is related to autonomic regulation of the cardiovascular system, the sympathetic and parasympathetic nervous system provides an excitatory or inhibitory stimulation and to the heart, respectively. [8 p.228]

Table 1. Effect of ANS on various organs

Organ	Sympathetic effect	Parasympathetic effect
Eye	Dilated pupil & slight relaxation of the ciliary muscle	Constricted pupil & ciliary muscle
Sweat glands	Copious sweating	Sweating on hand palms
Blood vessels	Constricted	No effect
Heart	Increased heart rate with an increased force of contraction, dilated or constricted coronaries	Decreased heart rate with a reduced force of contraction, dilated coronaries
Lungs	Dilated bronchi	Constricted bronchi
Liver	Glucose released	Slight glycogen synthesis
Systemic arterioles	Constricted abdominal viscera and skin	No effect

2.2 The Heart

The heart is the most important part of the human body, which receives deoxygenated blood from the body, oxygenated it via the lungs, and then pumps the oxygen and nutrition-rich blood through the whole body. At normal resting condition, the heart of an adult human pumps about 5 liters of blood in a minute. [8 p.378] It contains two double-chambered pumps, one pumps oxygen-poor blood through the lungs, and the other pumps oxygen-rich blood through the peripheral organs, and they know as the right and left heart, respectively. The chambers are named as the atrium (right atrium and left atrium) and the ventricle (right ventricle and left ventricle). The atrium works as an initial weak pump, and it pumps to flow the blood to the ventricle while the ventricle pumps the blood to the destination. [11 p.103]

The atria and ventricles are separated by a muscle wall, which prevents the mixture of blood. The heart has four valves to control the blood flow from the heart and within the heart. Two one-way atrioventricular (AV) valves control the blood flow from atria to ventricles, and two one-way semilunar valves control the blood flow from the ventricles to the lungs and peripheral organs. Blood enters the heart by two large veins, superior vena cava and inferior vena cava. [8 p.378-379] The heart is made by a special kind of muscle known as cardiac muscle. The cardiac muscle is different than the normal skeletal muscle due to its' ability of more prolonged contraction and higher conduction. There's another special type of muscle in the heart known as the excitatory and conductive muscle, which contract weakly regarding the cardiac muscle. The heart muscle fiber contains a latticework frame of interconnected heart muscle cells. When one cell is excited, due to the interconnection, the action potential spread to the other cells. [11 p.103-104]

The heartbeats occur in a rhythmic way, and the heart's rhythmicity is maintained by a special mechanism. Special excitatory and conduction muscle fibers maintain this rhythmicity by discharging action potential automatically in a rhythmic way and conducting it through the heart. The repeating contraction and relaxation phase of a heart is referred as the cardiac cycle. The cardiac cycle includes all the changes in the heart from the starting of a heartbeat to the beginning of the next heartbeat. [11 p.106] The contraction and relaxation phase of ventricles is known as systole and diastole, respectively. The atrial contraction occurs during the end of diastole, and relaxation occurs during the systole. During a cardiac cycle, the two atria contract simultaneously once they are filled with blood. The ventricles contract 0.1 to 0.2 *seconds* later after the atrial contraction. Ventricular contraction closes the AV valves and prevents the reverse flow of blood to the atrium. The semilunar valves are opened at the same time, and the blood left the ventricles towards the destination. The opposite things happen during ventricular relaxation. [8 p.381]

The sinus node (also known as the sinoatrial or SA node) plays a vital role in the cardiac cycle by generating an action potential simultaneously. Figure 2 represents the schematic of the heart with its' conducting system [37]. The conducting system of the heart includes the SA node, the AV node, the intermodal pathway between these two nodes, and the AV bundles spread from the AV node to the ventricles. SA node is situated in the superior posterolateral wall of the right atrium, and the generated impulse conducts through the whole heart. Cardiac action potentials are different than skeletal action potentials, and the major differences are long duration and speedy conduction. The initial spike in an action potential followed by a long plateau, which ensures long contraction time for the ventricles. Cardiac tissues are refractory to re-stimulation, i.e., only one action potential can stimulate the tissue at a time. [11 p.105-106, 117-118]

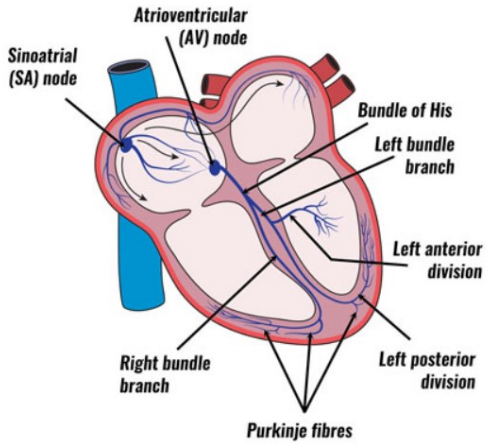

Figure 2. The heart and its' conducting system. The arrows are showing the conducting pathway of action potential through the heart.

The conduction of electrical impulses through the heart can be segmented into three portions. First, the generated action potential in the SA node spread through the atrial muscle fibers and reaches the AV node via the intermodal pathways. This conduction happens rapidly since the end of SA nodal fibers are connected to the atrial muscle fibers. Then, the AV node provides a slight delay before conducting the impulse to the

ventricles via AV bundles. This delay ensures that the atria have transferred the blood to the ventricles. Finally, the action potential arrived at the ventricles through the AV bundles and spread in the ventricular muscle fibers. The transmission from the AV node to ventricles is carried by a very high speedy fiber named as the Purkinje fiber. The AV bundle is divided into small branches and spread all over the ventricles so that once the action potential entered the AV bundle, it will spread through the entire ventricles immediately. This provides a simultaneous contraction of all portions of ventricles, which is essential for generating adequate ventricular pressure. Another important feature of the AV node is the one-way conduction to the ventricles, i.e., there's no way for the action potential to travel back to the atria. [11 p.117-121]

The heart is continuously innervated by the sympathetic and parasympathetic nerve fibers. Figure 3 shows the distribution of sympathetic and parasympathetic nerve fibers in the heart. [38] The parasympathetic nerve fibers, coming from the Vagus nerve, are mainly distributed near to the SA and AV node while the sympathetic nerve fibers are distributed all over the heart. [11 p.113] The cardiac control center is located in the medulla oblongata of the brain stem. Higher brain centers and the sensory feedback from the baroreceptors in aorta and cardiac arteries can affect the cardiac control center and finally, change in the heart rhythm. It's important to mention that this neural control from the brain is essential for maintaining different heart rhythms in different situations. [8 p.409]

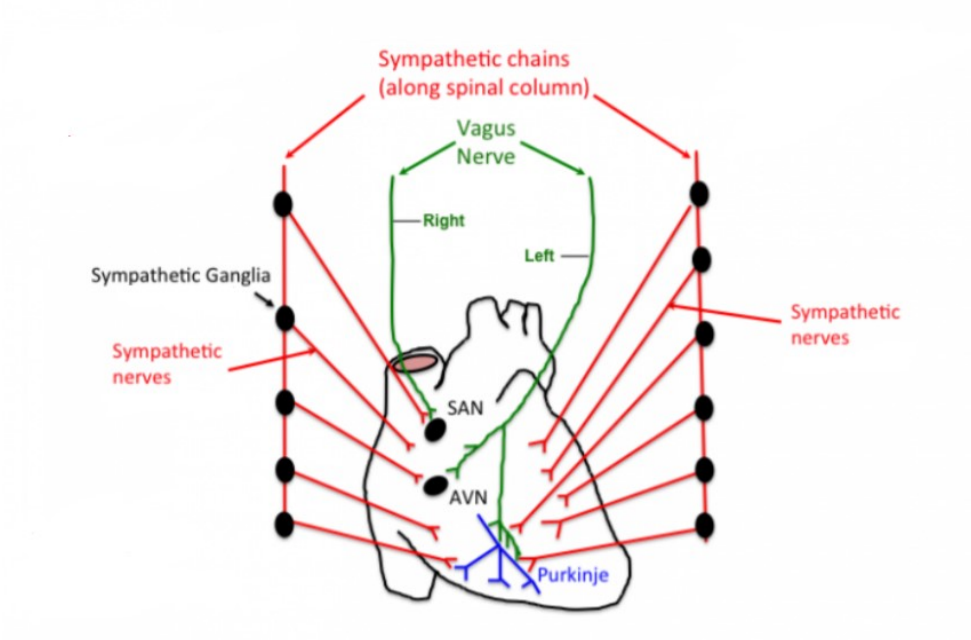

Figure 3. Distribution of sympathetic and parasympathetic nerve fibers in heart.

The SA node is known as the pacemaker of the heart since it generates a steady and repetitive signal and maintains the normal heart rhythm. It also provides necessary changes in the repetitive signal generation to retain the optimal heart rate in different situations according to the neural feedback from ANS. [12] The SA node receives dual innervations from both sympathetic and parasympathetic nervous systems and results in an antagonizing effect. When the SA node is innervated by sympathetic stimulation, it increases the rate of action potential discharge. Due to the sympathetic innervation,

the conduction rate in all portions of the heart is increased, and the contraction force in cardiac muscles (both is atria and ventricles) is also increased. [11 p.120] By increasing the overall activity of the heart, the sympathetic stimulation results in a higher heart rate and higher cardiac output. [10] The parasympathetic stimulation provides the opposite effects. First, it decreases the discharge rate of the SA node and the excitability of the intermodal pathways, which results in a delay in the arrival of impulse at the AV node and weaker atrial contraction force. Since the parasympathetic nerve fibers are not spread all over the heart, it has no significant effect on ventricular contraction. Eventually, the slow conduction of impulse will result in a delay in ventricular contraction, lower heart rate, and lower cardiac output. Normal parasympathetic stimulation can decrease the heart rate to half of the normal, while strong parasympathetic stimulation can completely shut down the impulse generation and conduction at SA node. [11 p.121]

2.3 The Electrocardiogram

The electrocardiogram (ECG) demonstrates the electrical activity of the heart. Our body is a good conductor of electricity [8 p.387], and when the electrical impulse passes through the heart, the electrical current is not confined only in the heart as a portion of the current spreads towards the body surface through the surrounding tissues. [11 p.123] This electrical current can be measured by placed electrodes on the body surface. ECG is the most known and used biomedical signal. [14 p.21] The first ECG recording on human were conducted by Augustus Waller in the 1880s. [15 p.411] It's important to mention that ECG is not the recording of action potentials; instead, it's a recording of the production and conduction of action potential in the heart. [8 p.387]

ECG provides a non-invasive way to gather information about cardiovascular activity, and it's recorded in various clinical applications. The recorded ECG waveshape in normal condition describes the cardiac muscle activity and the conduction system at that time. [15 p.412] Any diversion in cardiovascular activity from the normal condition can change the ECG waveshape, and that's why ECG is a major clinical tool for the diagnosis of different cardiovascular diseases. The abnormalities in cardiac muscle can be recognized by analyzing the contour of different ECG waveshapes from different leads. Another common and major cardiovascular problem is abnormalities in heart rhythm, generally known as arrhythmias. [11 p.147] Beat-to-beat analysis of ECG with the help of various signal processing algorithms provides a clear insight of different arrhythmias. Discoveries of different variability patterns of ECG make it a more suitable analyzing tool for arrhythmias. [15 p.412]

In a cardiac cycle, the depolarization and repolarization of millions of cardiac cells generate different action potentials. These potential changes result in electrical current in the intracellular fluids, which can be shown from an ECG measurement. ECG is not a cellular level measurement, and it represents the composite activity of those cardiac cells. [15 p.415] Like the other biomedical signals, the ECG carries the information of physiological events. One cardiac cycle contains three distinct waveshapes in ECG, the P wave, the QRS complex, and the T wave. Sometimes, the QRS complex is presented as three separate waves, the Q wave, the R wave, and the S wave. [14 p.234] Figure 4 shows the full ECG waveform in a complete cardiac cycle. [16]

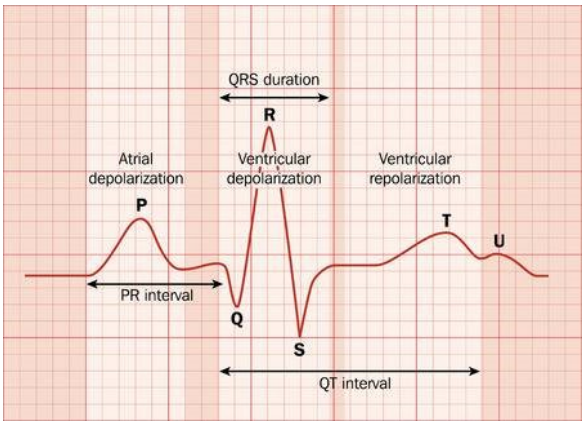

Figure 4. Full ECG waveform.

At the initial phase of a cardiac cycle, all cardiac cells are in the resting state, and the isoelectric line in the ECG represents this baseline situation. Before the atrial contraction, the SA node fires and results in the depolarization of atria. The P wave in ECG corresponds to this atrial depolarization. Since the atrial contraction is performed in a slow squeezing manner, and the atria are smaller with respect to the ventricles, the P wave appears as a slow waveform with low amplitude. It reaches the maximum when half of the atrial cells are depolarized. The duration of the P wave is around 80 *milliseconds*, and the amplitude is approximately 0.1 to 0.2 *millivolt*. [14 p.234] Once the atrial depolarization is finished, ECG returns to the baseline and remains there until the beginning of ventricular depolarization. Atrial repolarization doesn't produce any visible wave in the ECG since it's concurrent with the ventricular depolarization. [15 p.417] There's a slight delay before the ventricular contraction, and this is known as the PQ segment in the ECG. The duration of this PQ segment is about 80 *milliseconds*, and it's useful to recognize the baseline of the ECG signal. The sharp upward deflection in the ECG, known as the QRS complex in the ECG, is the result of rapid ventricular depolarization. Due to the higher mass of the ventricles, the QRS complex has a large amplitude with respect to the other waveshapes in ECG. This sharp wave has 1 *millivolt* amplitude with 80 *milliseconds* duration. Once the ventricles are entirely depolarized, the ECG returns to the baseline. After the QRS wave, an isoelectric segment of 100 to 120 *milliseconds* appears in the ECG. This segment is known as the ST segment, and the plateau portion of the action potential is responsible for this. Ventricular repolarization is represented by the slow and elusive T wave of the ECG signal. T wave has a slow amplitude of 0.1 to 0.3 *milliseconds* and a duration of 120 to 60 *milliseconds* approximately. [14 p.234-235]

The ECG is usually measured by strapping several electrodes on the body surface. The voltage difference between two electrodes in ECG recording is known as the lead. There are two types of lead, the unipolar lead, which represents the variation of voltage of a single electrode with respect to a reference electrode, and the bipolar lead which represents the voltage difference between two electrodes. The reference electrode in the unipolar lead is placed in such a position so that it can keep a constant voltage during the cardiac cycle. [15 p.419]

The standard 12-lead ECG is the most common and popular ECG recording system for clinical purposes. It is the combination of three different lead configurations, three bipolar limb leads (I, II & III), three augmented unipolar limb leads (aVR, aVL & aVF), and six unipolar precordial leads (V1, V2, V3, V4, V5 & V6). The bipolar leads

I, II & III, represent the voltage difference between right arm to left arm, right arm to left leg, left arm to left leg respectively, and the right leg is used as the ground lead. The unipolar leads aVR, aVL & aVF are placed on the right arm, left arm, and left foot, respectively, and the other six unipolar lead (labeled as V1 to V6) are placed on the chest. [8 p.388]

2.4 Heart Rate Variability

Over the years, heart rate variability (HRV) has earned a widespread research interest due to its ability to assess autonomic nervous system activity. There is a notable relationship between ANS and cardiovascular mortalities, and a proper assessment of ANS activity is really important. [15 p.567-568] HRV provides a noninvasive way of investigating and assessing the ANS state. Our heart rate is not constant since the SA node maintains a healthy irregularity in the heart rate. The heart rate is directly controlled by the ANS, and the sympathetic and parasympathetic nervous system interacts with the heart simultaneously. The origin of HRV lies within this simultaneous interaction. HRV indicates both the variations in the interval of consecutive heartbeat (R-R interval) and the variations in the instantaneous heart rate. [17] Though the simple derivation of HRV makes it a powerful tool for both research and clinical purpose, it's crucial to ensure that HRV is analyzed correctly.

Before explaining any details about HRV, it's important to know the clinical relevance of it. Since HRV is a tool for analyzing ANS imbalances, any health disorder, i.e., cardiovascular diseases, can create a significant change in HRV. Regular monitoring of HRV can help a person to lead a healthier life. Numerous research works have performed over the last few decades to find out the exact relationship between HRV and different cardiovascular diseases. During normal sinus rhythm, a high HRV indicates better cardiovascular fitness and proper ANS adaptability, while a low HRV refers to the opposite. [18] Abnormal HRV is used as a risk marker for both arrhythmic and non-arrhythmic cardiac events. In different studies, it's shown that HRV is a predictor of sudden arrhythmic death. [19, 20, 21] Patients suffering from congestive heart failure or myocardial infarction (MI) showed a weakened HRV. [22, 23]

In recent years, due to the development of wearable devices, it becomes easy to collect and analyze long term HRV recordings. Many recent studies try to show the relationship between HRV and stress (both mental and physiological). [24, 25, 26, 27, 28] Melillo P et al. [24] showed that the analysis of short-term HRV could determine the stress automatically in real-life scenarios. In that study, the participants were participating in university examinations. Salahuddin L. et al. [25] used the Stroop color-word test to determine the stress by analyzing short term HRV. Hynynen E. et al. [29] used HRV from sleep and orthostatic test to categorize the participants in low and high stress group.

Studies performed on the data from professional and elite athletes to recognize the state of physiological stress and recovery. Proper assessment of stress and recovery can reduce the occurrence of injury during exercise. Plews D. J. et al. [30] described the relationship between HRV and the training adaptation in elite endurance athletes. The described method was using HRV as a day-to-day analysis tool to assess the training status and readiness for training. They also focused on the unique individual HRV of the elite endurance athletes. It's important to note that elite athletes have

different HRV characteristics than normal people due to their extraordinary capability. HRV as a parameter of assessing physiological stress and recovery from stress is a relatively new research topic, and hopefully, in the future, we will have some more exciting results.

During the interpretation of HRV, it's important to take account of the other information of the participants, i.e., age, gender, etc. The influence of gender in HRV is dependent on age and measurement, and HRV in healthy subjects declined with aging. [31] Same subject can show different HRV in the same condition in a different measurement time of the day. [32] Change in food habits, inadequate sleep, high working load, drugs, smoking can also affect the HRV.

2.5 System Identification Procedure

System identification is a wide concept that covers the whole process of building mathematical models from the observed data of a dynamic system. It has vast applications in various fields of science since our surroundings are full of numerous dynamic systems. In general words, system identification refers to the procedure of building a model by estimating the system parameters from the available prior knowledge and input-output data from the real system. [33 p.1, 34 p.1]

A system is an object which can produce an observable signal from the interaction of different kinds of variables. The observable signal is known as output, and it can be manipulated by external signals. If the user can control the external signal, then it's called input; otherwise, it's a disturbance of the system. The disturbance is either measurable or unmeasurable, and if it's unmeasurable, it's usually noted as system noise. If the output of a system depends on not only the current input values but also the past system behaviors, it's called a dynamic system. [34 p.1-2]

A solar-heated house is a good example of a dynamic system with all system components. The working principle of this system is such that the sun heats the air of the solar panel, which is fanned to the heat storage where the pebbles stored the heat, and finally, the heat is transmitted to the house. A mathematical model can be designed for this system to observe how the radiation from the sun and speed of the fan affect the heat storage temperature. In this system, the speed of the fan is the input variable, solar radiation is the measured disturbance, and the temperature of the heat is the output variable. [33 p.2]

A mathematical model of a real system is always based on approximation. The real system can have various complexities like insufficient observation data, limited prior knowledge, etc. These complexities diminish the possibility of building an exact model of the real system. However, if it's possible to build the exact model, the model will be too complex to use. [34 p.5] The system identification procedure includes different iterative steps to find the perfect approximation model of the real system. [33 p.13]

The system identification procedure has three major entities; data, prior knowledge, and objectives. These entities are depended on each other. Prior knowledge and objectives define the system identification task. Often the data is collected based on the available prior knowledge and objectives. In the same way, the collected data can have an impact on other entities and forced to adjust them in a meaningful way. It's worth to mention that the prior knowledge and objectives are the important parameters to determine all other system identification procedures.

The selection of model set may be the most challenging task of the system identification procedure. Based on the model structure, there are three major types of model, white-box model, grey-box model, and black-box model. A white-box model is deterministic in nature, and the model is based on the physical laws and relationship between physical parameters. This model doesn't provide any uncertainty, and it's over-specified. Complete prior knowledge of the real system is an essential requirement for designing a white-box model. In the real world, it's often happened that some physical parameters are uncertain or not familiar. These parameters are needed to be estimated from the observed data, and this type of model is called the grey-box model. Finally, in black-box modeling, the primary aim of the model is to fit the observed data, and it doesn't depend on the physical laws of the process. [33 p 13-15, 34 p.11-12] Besides the structure, there are several ways of describing a model. Transfer function equation, differential equation, state-space equation, and impulse response are usually used for model representation. A model can be represented either in discrete time or in continuous time. [35]

In the system identification procedure, the criterion function is defined, which measures the fit between the model output and the observed data. The identification method determines how well the designed model can reproduce the measured data. The last phase of the system identification procedure is the validation of the designed model. Model validation determines whether the designed model is appropriately fulfilling its' purposes and planned usages. Though the parameter estimation process will give us a suitable model structure, it's not accurate to assume that the model is appropriate to fulfill the necessary purposes. Model validation covers many general aspects like how well the model serves the desired goals, how good the model is to define the true system, how well the model fits with the real-time data etc. If the model is not good enough, then the previous procedures are repeated until the desired model is found. [34 p.11-12] This is the situation where the system identification procedure enters the system identification loop and continue iterating. It's wise to remember that, sometimes the system identification loop can fail to provide the best outcome also. This can happen for various reasons, i.e., bad choice of criterion function, lack of information in observed data, inappropriate choice of model set, and many more. [33 p.13-15]

3. METHODS

This chapter provides the details explanation of the dataset and different methodologies that are used in this thesis. Figure 5 shows the overall working procedure of the thesis, including the implemented methodologies. Before doing any analysis, it's necessary to perform a few pre-processing steps to the raw ECG signal. The preprocessing steps are omitted from the general pipeline since these steps are performed in the R peak detection. HRV signal is created from the processed ECG signal, and finally, the constructed HRV signal is segmented into different training and validation signals, which are used for model training and validation, respectively.

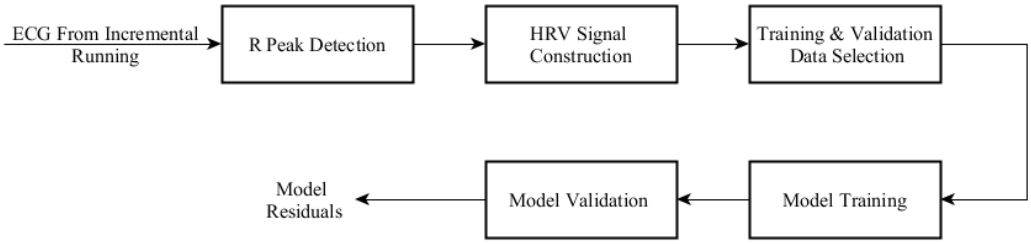

Figure 5. Block diagram of the general working procedure.

3.1 Dataset & Protocols

The dataset contains data from 10 healthy persons, and the subjects were performing a maximal running test on a treadmill. The general characteristics of the participants are given in table 2. The subjects participated in the study voluntarily. They were fully aware of the data collection protocol and the purpose of the study, and they gave their written consent. All of them were familiar with exercise testing. The blood pressure level and ECG pattern were normal for the participants, which were checked by a practitioner before the test. There were no complaints from the participants regarding the data collection process.

Table 2. General characteristics of the subjects.

Characteristics	Mean	STD	Min	Max
Age ($years$)	38.25	5.34	25	50
Weight (Kg)	69.37	12.57	46.6	103.9
Height (cm)	175.13	8.8	160	200
VO2Max (ml/min)	53.09	5.66	39.8	65.20

Few general instructions were given to the participants beforehand. They were suggested to wear comfortable clothes for the test and avoid food and drinks prior to the test. They were also recommended to avoid exercise and any other highly intensive physical activity on the test day. The maximal running test was performed in a controlled environment. For ECG measurement, a three-channel Faros 360 chest strap with a sampling rate of $1\ kHz$ was used. This device also provided three-channel accelerometer signals at $1\ kHz$. Additionally, a spirometer was used to record oxygen consumption and respiratory frequency.

At the beginning of the data collection protocol, the subjects had a resting segment (stage I). During the resting period, the participants stood quietly without any massive body movement, and the baseline cardiovascular measurements were measured at that time. The treadmill was switched on from stage II with a starting speed of eight *km/h* for males and seven *km/h* for females. Each stage lasted for three minutes, and there was a pause of 15 to 30 *seconds* between two successive stages. The blood lactate measurement from the fingertip was taken in the break. The treadmill speed was increased by one *km/h* in every stage. The schematic of the maximal running test is shown in figure 6.

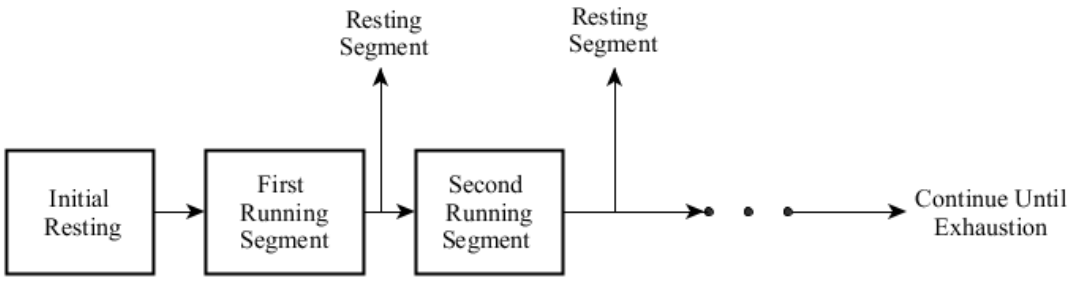

Figure 6. Block diagram of incremental running procedure.

3.2 R Peak Detection & HRV Signal Construction

The R peaks of the ECG signal are detected by using the Pan-Tompkins algorithm. [36] This algorithm analyzes the amplitude, slope, and width of the QRS complexes of the ECG signal. Often raw ECG signal is affected by different artifacts like power line interference, baseline wandering, etc. The Pan-Tompkins algorithm provides a series of filters to remove those artifacts. The block diagram of the Pan-Tompkins algorithm is illustrated in figure 7.

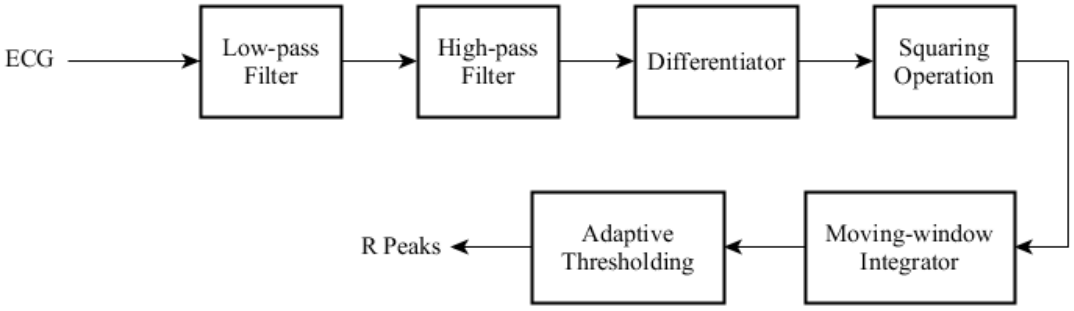

Figure 7. Block diagram of the Pan-Tompkins algorithm.

The combination of low-pass and high-pass filter works as a bandpass filter. The lowpass filter is used to remove the power line interference and other high-frequency noises of ECG. This filter has the following transfer function,

$$H(z) = \frac{1}{32}\frac{(1 - z^{-6})^2}{(1 - z^{-1})^2} \tag{1}$$

The input-output relationship of the filter is given below,

$$y(n) = 2y(n-1) - y(n-2) + \frac{1}{32}[x(n) - 2x(n-6) + x(n-12)] \tag{2}$$

The high-pass filter is introduced to reduce the effect of baseline wandering and muscle noises. The following two equations describe the transfer function and input-output relationship of this filter.

$$H(z) = z^{-16} - \frac{1}{32}\frac{(1 - z^{-32})}{(1 - z^{-1})} \tag{3}$$

$$y(n) = y(n-1) - \frac{1}{32}x(n) + x(n-16) - x(n-17) + \frac{1}{32}x(n-32) \tag{4}$$

In the time domain, the derivative operation works in such a way that the output is zero for constant input and high for any large change in input. Thus, the derivative operation provides a large gain in the high-frequency components of the QRS complex as well as restrains the low-frequency components of P and T wave. This operation is specified as,

$$y(n) = \frac{1}{8}[2x(n) + x(n-1) - x(n-3) - 2x(n-4)] \tag{5}$$

Point to point squaring provides non-linear amplification on the large changes that arise from the QRS complex. The output result becomes positive, and the minor differences from the P and T wave are minimized. The squaring equation is given below,

$$y(n) = [x(n)]^2 \tag{6}$$

The final preprocessing step is the smoothing of the squared signal by using a moving-window integrator. Smoothing is needed since the derivative-based operation gives more than one peak in one QRS complex duration. Following equation is used for the smoothing operation,

$$y(n) = \frac{1}{N}[x\{n-(N-1)\} + x\{n-(N-2)\} + \cdots + x(n)] \tag{7}$$

where N is the number of samples for the moving-window integrator. N should be selected precisely since too large N can combine the QRS complex and T wave, and too small N can create multiple peaks for a single QRS complex. With a $200\ Hz$ sampling rate, $N = 30$ is enough, and the integrator has a moving window of $150\ ms$ which is approximately equal to the widest possible QRS complex. The pre-processing steps on the ECG signal are illustrated in Figure 8.

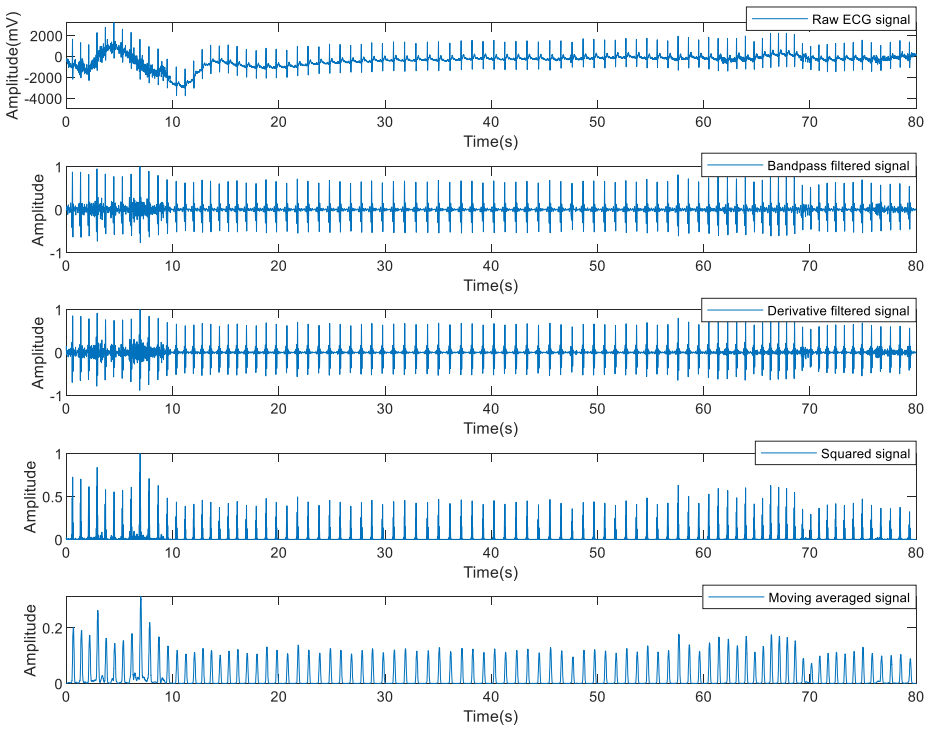

Figure 8. Pre-processing steps in the Pan-Tompkins algorithm. The top-most sub-figure shows the raw ECG signal, and the following four sub-figures illustrate the output of the bandpass filter, derivative-based filter, squared operation, and moving average filter respectively.

The moving-window integrator provides a pulse-shaped output. Adaptive thresholding is used to determine whether a pulse belongs to the QRS complex or not. The adaptive thresholding procedure keeps the estimation of current noise and signal

peaks. In this way, the Pan-Tompkins algorithm adapts with the changes in the ECG signal. A peak is the local maxima within a specific time interval, which is detected by spotting the change of direction of the output signal. The time interval is selected in such a way that only one QRS complex can occur within that interval. From the physiological point of view, the minimum distance between successive R-R peaks should be $200\ ms$, and hence, it is selected as the time interval for peak detection. This time interval eliminates the possibility of multiple peak detection for a single QRS complex.

After detecting a peak, two threshold values, THR_SIG and THR_NOISE are used to classify the peak as R peak or noise by using the following equations.

$$PEAK > THR\_SIG \ \rightarrow PEAK = R\_PEAK \tag{8}$$

$$THR\_NOISE < PEAK < THR\_SIG \ \rightarrow PEAK = NOISE\_PEAK \tag{9}$$

where $PEAK$ is the detected peak. The threshold values are calculated from,

$$THR\_SIG = NOISE\_LEVEL + 0.25(SIG\_LEVEL - NOISE\_LEVEL) \tag{10}$$

$$THR\_NOISE = 0.5 * THR\_SIG \tag{11}$$

where SIG_LEVEL and $NOISE\_LEVEL$ represent the estimation of current signal and noise peak level. After each peak detection, these two values are updated by using the following equations,

$$SIG\_LEVEL = 0.125 * PEAK + 0.875 * SIG\_LEVEL \tag{12}$$

$$NOISE\_LEVEL = 0.125 * PEAK + 0.875 * NOISE\_LEVEL \tag{13}$$

If $PEAK$ is classified as an R peak, the SIG_LEVEL will be updated; otherwise $NOISE\_LEVEL$ will be updated. To initialize these two values, approximately $2\ s$ of the training phase is required at the beginning. After every peak detection (R peak or noise), the threshold values are also modified by using the updated SIG_LEVEL and $NOISE\_LEVEL$. The adjustment of threshold values is the key feature of the Pan-Tompkins algorithm.

To improve the performance of R peak detection, the Pan-Tompkins algorithm includes a search back procedure. If no R peak is detected by THR_SIG and equation (8) within a long-time period, the algorithm provides a search back option. For the search back operation, the algorithm stores two average R-R interval values, RR_AVG_1 is the average of eight latest beats and RR_AVG_2 is the average of eight latest beats within a specific limit, from RR_LOW_LIMIT to RR_HIGH_LIMIT. The limit values are defined by,

$$RR\_LOW\_LIMIT = 0.92 * RR\_AVG\_2 \tag{14}$$

$$RR\_HIGH\_LIMIT = 1.16 * RR\_AVG\_2 \tag{15}$$

The time interval to trigger a search back is RR_MISSED_LIMIT, which is defined as,

$$RR\_MISSED\_LIMIT = 1.66 * RR\_AVG\_2 \quad (16)$$

During the search back, the second threshold value, THR_NOISE is used to detect the R peaks. Equations (8) and (12) are modified with the following changes,

$$PEAK > THR\_NOISE \rightarrow PEAK = R\_PEAK \quad (17)$$

$$SIG\_LEVEL = 0.25 * PEAK + 0.75 * SIG\_LEVEL \quad (18)$$

In the final stage, the algorithm provides a T wave identification if the R-R interval is less than $360\ ms$. By observing the slope of the detected peak, it's classified as either an authentic R peak of the QRS complex or an abnormal T peak of the T wave. If the slope is lower than 50% of the previously detected R peak, then it's considered as noise (T peak). Figure 9 shows the selected R-peaks from the same ECG signal of figure 8 after using adaptive thresholding, search back procedure, and T-wave identification.

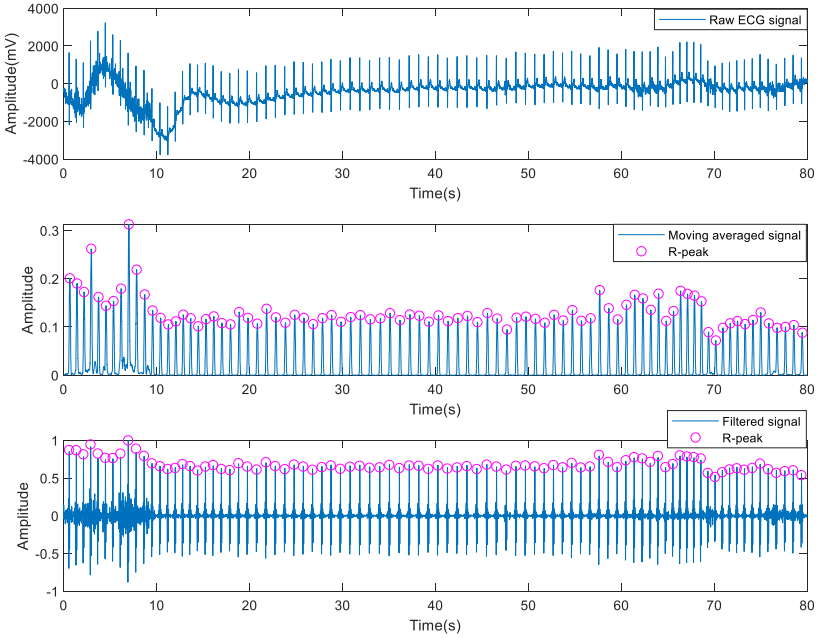

Figure 9. Detected R-peaks by using the Pan-Tomkins algorithm. The top-most sub-figure shows the raw ECG signal, the next one shows the location of R-peaks in the moving averaged signal (the output from the pre-processing step) and the last sub-figure shows the same R-peaks in the filtered ECG signal.

After detecting the R-R intervals, the HRV signal is constructed by interpolating the R-R intervals at $8Hz$, and the ectopic beats of the HRV signal are interpolated.

3.3 Polynomial Model

In this section, we will try to describe the transfer function of a model and the general representation of a system. Figure 10 shows a single input single output (SISO) linear time-invariant (LTI) system with an additive noise where u(t), y(t), and v(t) is the system input, output, and disturbance respectively.

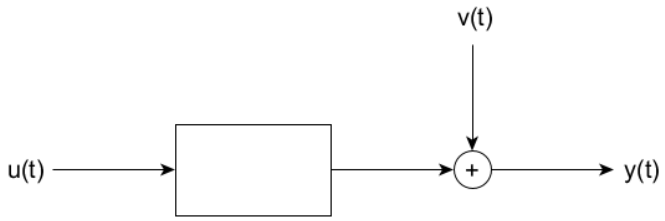

Figure 10. LTI system with additive noise.

To describe the system, the convolutional integral and convolutional sum is used for continuous-time representation and discrete-time representation, respectively. [34 p.19] The discrete-time representation of this system is given below.

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) + v(t) \tag{19}$$

where, $t = 0,1,2,3 \ldots \ldots$ represents the sampling instances and $g(k)$ specifies the discrete impulse response of the system.

Though it's impossible to know the value disturbance in ahead, previous disturbance values can help to make a guess for the upcoming ones. $v(t)$ can be assumed by using the same approach of equation (19),

$$v(t) = \sum_{k=0}^{\infty} h(k)e(t-k) \tag{20}$$

where the term $e(t)$ belongs to a series of independent random variables, which has a mean and variance of 0 and λ, respectively. So, equation (19) becomes,

$$y(t) = \sum_{k=1}^{\infty} g(k)u(t-k) + \sum_{k=0}^{\infty} h(k)e(t-k) \tag{21}$$

After introducing the backward time shift operator q^{-1} in equation (21),

$$y(t) = \sum_{k=1}^{\infty} g(k)q^{-k}u(t) + \sum_{k=0}^{\infty} h(k)q^{-k}e(t) \tag{22}$$

The backward shift operator works in such a way that,

$$q^{-1}u(t) = u(t-1) \tag{23}$$

Now, equation (22) can be rewritten as,

$$y(t) = G(q)u(t) + H(q)e(t) \tag{24}$$

where, $G(q)$ and $H(q)$ are defined by the following equations,

$$G(q) = \sum_{k=1}^{\infty} g(k)q^{-k} \tag{25}$$

$$H(q) = \sum_{k=0}^{\infty} h(k)q^{-k} \tag{26}$$

Equation (24) is the basic representation of a linear system where $G(q)$ is known as the transfer function of the system.

In dynamic systems, the basic input-output relationship can be described by using the linear difference equation, which is given below,

$$\begin{aligned} y(t) + a_1 y(t-1) + a_2 y(t-2) + \cdots + a_{n_a} y(t-n_a) \\ = b_1 u(t-1) + b_2 u(t-2) + \cdots + b_{n_b} u(t-n_b) + e(t) \end{aligned} \tag{27}$$

where, e(t) is the white noise of the system. This equation represents the common form of the polynomial model structure since all polynomial models can be obtained by modifying this equation. After introducing backward shift operator and transfer functions, equation (27) becomes,

$$A(q)y(t) = B(q)u(t) + e(t) \tag{28}$$

where $A(q)$ and $B(q)$ are defined as

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \cdots + a_{n_a} q^{-n_a} \tag{29}$$

$$B(q) = b_1 q^{-1} + b_2 q^{-2} + \cdots + b_{n_b} q^{-n_b} \tag{30}$$

The coefficients from the above two equations will be estimated through the parameter estimation process. The parameter vector is defined as,

$$\theta = \left[a_1, a_2, \ldots, a_{n_a}, b_1, b_2, \ldots, b_{n_b}\right]^T \tag{31}$$

In equation (28), $A(q)y(t)$ is the autoregressive part and $B(q)u(t)$ is the exogenous part, and it can be rewritten as,

$$y(t) = \frac{B(q)}{A(q)}u(t) + \frac{1}{A(q)}e(t) \tag{32}$$

Equation (32) is the general equation of the autoregressive model with exogenous input (ARX).

3.4 Parameter Estimation for the ARX Model

ARX model structure is the simplest continuous-time representation of a polynomial model. Figure 11 shows the block diagram of the ARX model structure. From equations (26) and (32), it's visible that the transfer function of the ARX model is a ratio of two polynomials, $B(q)$ and $A(q)$ hence, known as a rational transfer function. [34 p.114] The noise model is represented as $1/A(q)$, and it's integrated into the system dynamics. For this reason, it's not possible to estimate the noise model separately in the ARX model. [35]

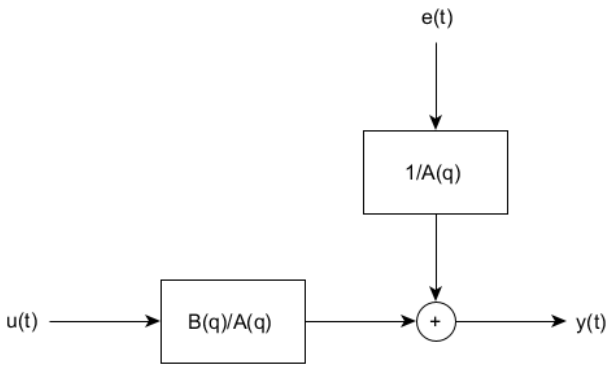

Figure 11. Structure of ARX model.

The parameter estimation process includes representing the ARX model structure in the form of a linear regression model and then solve it to find the value of unknown parameters by using least square estimation. Linear regression is a commonly used statistical tool to explain the relationship between one observer variable and one or more explanatory variables, which are also known as the regressors. [34 p.61] The linear relationship between observer variable $y(t)$ and p regressors is given by,

$$y(t) = \phi_1(t)\theta_1 + \cdots + \phi_p(t)\theta_p + e(t) \tag{33}$$

where, $t = 1,2, \dots, N$, $e(t)$ is the unobserved error and $\phi_1 \dots \phi_p$ are the regressors. The vector-matrix form of equation (33) is given by,

$$y = \Phi\theta + e \tag{34}$$

where, Φ is a $N \times p$ regression matrix and the vectors y, θ and e are defined as,

$$y = [y(1), y(2), \dots, y(N)]^T \tag{35}$$

$$\theta = [\theta(1), \theta(2), \dots, \theta(N)]^T \tag{36}$$

$$e = [e(1), e(2), \dots e(N)]^T \tag{37}$$

The output of the ARX model can be represented as the linear regression form described in equation (33). The predicted ARX model output is given by,

$$\hat{y}(t, \theta) = \phi(t)^T \theta \tag{38}$$

where, $t = 1,2, \dots, N$, the parameter vector θ is defined in equation (31), and the data vector $\phi(t)$ is defined as,

$$\phi(t)^T = \left[-y(t-1), -y(t-2), \dots, -y(t-n_a), u(t-1), u(t-2), \dots, u(t-n_{b)}\right] \tag{39}$$

The ARX model structure assumes the noise as a white noise, and the predictor equation doesn't depend on the system noise. [33 p.82] That's why the noise term is absent in equation (38). The predictor equation of ARX model is given below,

$$\hat{y}(t, \theta) = B(q)u(t) + [1 - A(q)]y(t) \tag{40}$$

It is worth to mention; this prediction equation is used when the system noise is either insignificant or difficult to guess. [33 p.82]

The predicted ARX model output can be represented in a vector-matrix form like the equation (34) by defining the output vector and the regressor matrix. If N is greater than the maximum of n_a and n_b, the output vector y and regressor matrix Φ is defined as,

$$y = [y(\max(n_a, n_b), \dots, y(N)] \tag{41}$$

$$\Phi = \begin{bmatrix} -y(n_a - 1) & \cdots & -y(0) & u(n_a - 1) & \dots & u(n_a - n_b) \\ -y(n_a) & & -y(1) & u(n_a) & & u(n_a - n_b + 1) \\ -y(n_a + 1) & \vdots & \vdots & u(n_a + 1) & \vdots & \vdots \\ \vdots & & & \vdots & & \\ -y(N-1) & \cdots & -y(N - n_a) & u(N-1) & \dots & u(N - n_b) \end{bmatrix} \tag{42}$$

This regressor matrix is valid for $n_a \geq n_b$ and similar regressor matrix can be produced for $n_b > n_a$ also. [34 p.118] By knowing the output value $y(0), y(1), \dots, y(N)$ which are corresponded to the input values $u(0), u(1), \dots, u(N)$, the elements of the parameter vector can be estimated. The final equation for parameter estimation becomes,

$$y = \Phi\theta \tag{43}$$

The least-square method estimates the parameter values by minimizing the prediction error. From equation (38), the prediction error of ARX model structure id given by,

$$\varepsilon(t) = y(t) - \hat{y}(t, \theta) = y(t) - \phi(t)^T \theta \tag{44}$$

The unknown parameter vector θ should be defined in such a way that it will provide a minimum overall squared error. The least-square method introduces a scalar function to minimize the error. This function is also known as the least-square objective function, [34 p.62] which is stated below,

$$J(\theta) = \sum_{t=1}^{N} \varepsilon^2(t) \tag{45}$$

Equation (45) can be represented in matrix notation, which is,

$$J(\theta) = \varepsilon^T \varepsilon \tag{46}$$

After combining the equations (43), (45) and (46),

$$J(\theta) = (y^T - \Phi^T \theta^T)(y - \Phi\theta) \tag{47}$$

The minimal value of this scalar function can be obtained if the gradient of J with respect of θ is zero. This condition will be true if and only if,

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T y \tag{48}$$

Here, $\hat{\theta}$ is the estimated parameter vector, which will provide the minimum squared error. The mathematical calculations between the equations (47) and (48) are skipped in here. The detailed explanation of this solution can be found in "Appendix II of [33]" and "Appendix A of [34]".

The ARX model structure can be defined by three parameters, $ARX = (n_a, n_b, n_k)$ where, n_a defines the polynomial order of $A(q)$, n_b defines the polynomial order of $B(q)$ and n_k states the input-output delay.

The fit percentage of a designed model is calculated by using the following equation,

$$Best\ fit = \left(1 - \frac{|y - \hat{y}|}{|y - \bar{y}|}\right) \times 100 \tag{49}$$

where, y is the original output, $\hat{y}$ is the predicted output, $\bar{y}$ is the mean of the original output. A fit percentage value of 100% indicates the perfect fit while 0% indicates that the output is guessed as a constant ($\hat{y} = \bar{y}$). The fit percentage value can be negative also, which means that there might be something wrong with the parameter estimation of the model. [35]

3.5 Modeling Approach

Since we are focusing on predicting the future R-R intervals by using the history of previous R-R intervals, a simple ARX model is a good choice for that as ARX is the simplest form of the polynomial model. Initially, we tried to see the effect of different

model orders and find out which model orders are not suitable to predict the R-R intervals. In the previous chapter, we discussed the ARX model structure and introduced the parameters of ARX model n_a, n_b, and n_k. To build our model and define parameters, we used two different approaches, the ARX model with input and output signal and the AR model with a single output without any input. We preferred the ARX model rather than the AR model since we have a hypothesis for adding additional inputs for our prediction model. We continued with both modeling approaches so that later we can compare their performances.

In the first approach, we chose the HRV signal as the model output and the same signal with a shift of one to eight sample points as the input signal. This means the output $y(t)$ at a given time t corresponds to the input $x(t+n)$ where $n = 1,2 \dots 8$, and it represents the future value. We produced 8 input signals for the single output signal and compare how the model performs with the increasing input-output sample difference. By choosing the parameters n_a and n_b, we can predict the output $y(t)$ from the input and previous output values. If we choose $n_b = 1$, the output will depend on the current input and n_a previous outputs. Since the HRV is interpolated at $8Hz$, the input-output difference of 8 samples means we are giving a future HRV sample, which is $1s$ after the current output as an input and predicting the current HRV signal based on the input and previous HRV signal. There's no time delay between the input and output, and the input signal affects the output immediately, there's no dead time in the model either, the third parameter n_k will be zero. Figure 12 shows a sample input-output signal for the ARX model. We are numbering the different inputs for the ARX model according to the input-output sample difference, i.e., $1st$ input means the output is shifted by one sample point from the input signal.

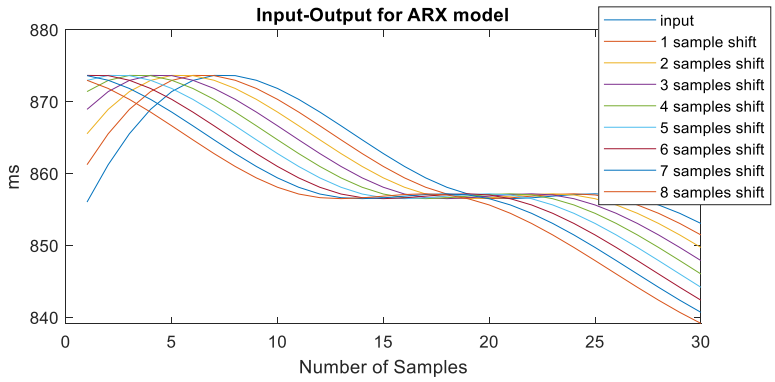

Figure 12. Input-output for the ARX model. The input and output correspond to the same signal, and the input is shifted by one to eight sample points.

If we remove the input signal from the ARX model, it will become the AR model. As our next approach, we used the HRV signal as the output of a time-series AR model. Since there is no input signal, the parameters n_b and n_k are not required for this model. We need to determine the parameter n_a and the present output $y(t)$ will depend on the n_a previous outputs. We can use the output signal from figure 12 for the AR model and design the model by ignoring the parameter n_b and n_k. To avoid any kind of confusion, the output signal for the AR model is shown in figure 13.

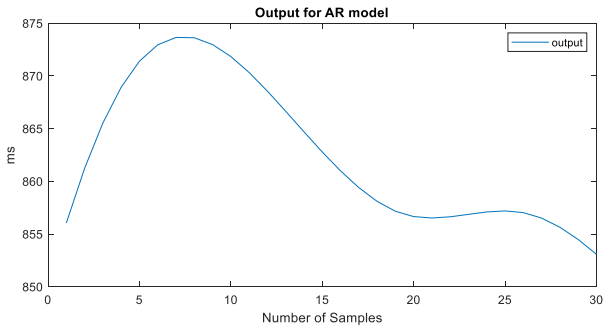

Figure 13. Output for the AR model.

Both of our approaches are predicting the current output based on the previous outputs. In addition to that, in the ARX model, we are giving a sample point form the near future as an input. We assume that by providing this input signal, our model will be more stable and provide increased performance with respect to its' AR counterpart. Through this study, we will see that how giving an input from the future helps our model to perform better.

4. EXPERIMENTAL RESULTS

This chapter contains the details about the used dataset for the experiment as well as the obtained results by following the procedures from the previous chapter.

4.1. Effect of Model Order

The first and foremost task of system identification is to find a suitable model and model order. After defining the modeling approaches, it's time to see the effect of model orders and choose the suitable one. For this purpose, we selected the same training data for different model orders and then checked the performance of each model with the same validation data. The training and validation data are shown in figure 14. For the first test procedure, we used the HRV signal from the normal resting condition. The HRV signal is constructed from the R-R intervals, and from now on, we are mentioning the word HRV rather than R-R intervals to avoid confusion.

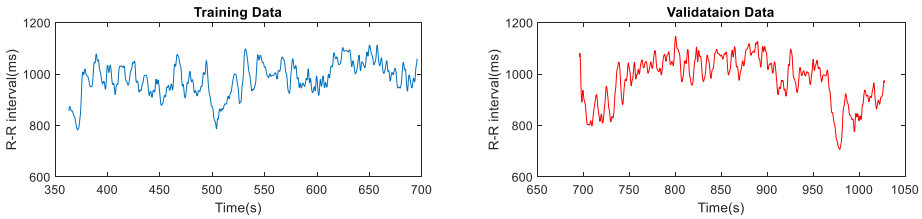

Figure 14. Training and validation data to observe the effect of different model orders. The training and validation data correspond to participant ID 1 and represent the HRV signal in normal condition (resting segment before the first running segment).

The performance of different models with the same training and validation data is shown in figure 15. The sub-figures in figure 15 stands for the two different approaches (ARX and AR). For the ARX model, we produced 8 input signals, which are discussed in the previous chapter, and the model performance with respect to each output is shown in the next figure.

From figure 15, we can say that if we use the HRV from normal condition to train and validate the model, the model performance for both ARX and AR do not depend on model order. After a few initial models, we have a pretty good fit percentage for both cases. For the ARX model the parameters n_b and n_k were set to 1 and 0 respectively, and for the AR model, both were ignored.

If we look at figure 14 again, the training and validation data is ~ 350 seconds long. This long training data is a reason for having good model performance even in the higher model orders. The training and validation data correspond to the normal resting condition's HRV signal, and there are no unusual changes in those signals. This also implies that we should have good fit percentages since the HRV dynamics remain the same for the training and validation data. So, we can say that it's hard to determine the perfect model order by using this training and validation data. We need to include the

running segment's HRV signal to compare the performance of different model orders. In the next experiment, we changed the validation data from figure 14 and replaced it by the HRV signal form the first running segment. Figure 16 demonstrates the training and validation data for the next experiment.

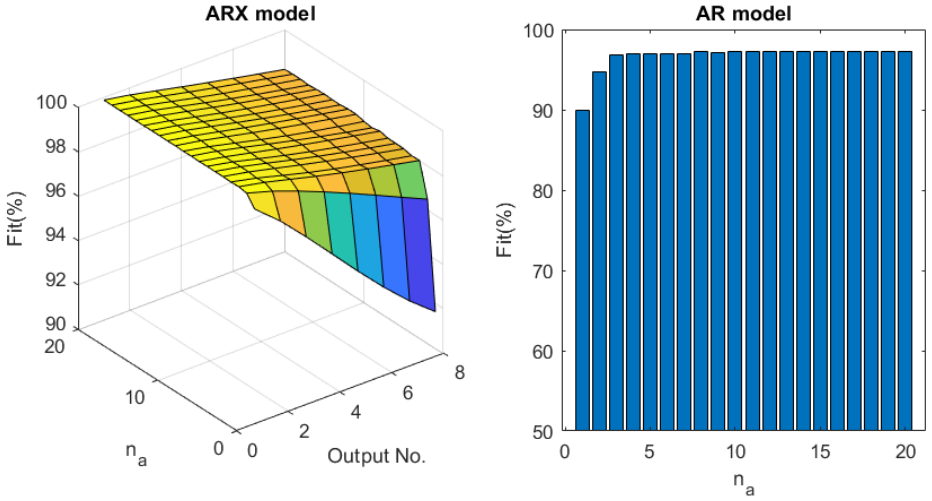

Figure 15. Performance of different ARX (left sub-figure) and AR (right sub-figure) models with the same training and validation data. In both cases, the parameter n_a is varying from 1 to 20.

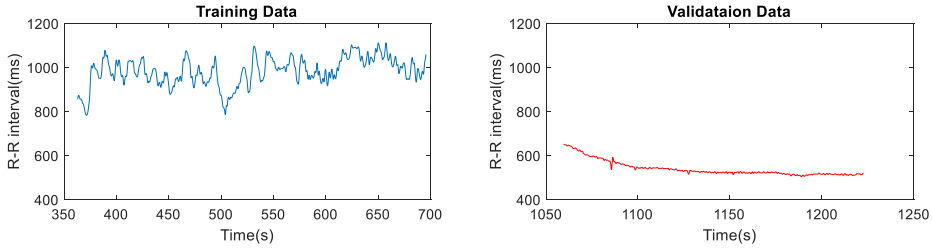

Figure 16. Training and validation data to observe the effect of different model orders. The training data is the same as the training data in figure 14, while the validation data corresponds to the HRV signal during the first running segment from participant ID 1.

The performance of different model orders with the different validation data is shown in the next figure. The validation data from figure 16 differs widely from 14, and we can see the impact of that difference in the model performances in figure 17. It's expected that the model performance will decrease if we use the training data from the resting segment to train the model and then validate it by using validation data from the running segment.

If we compare the left sub-figures (ARX model) of figure 15 and figure 17, a significant reduction in the overall fit percentage is seen. The lowest fit percentage is ~90% for the first one while it's ~60% in the latter one. We haven't observed any significant importance of the used sample shift in the input signal in figure 15. This scenario is changed substantially in the next figure since the increasing sample shift decreases the fit percentage gradually. If we compare the right sub-figures (AR model)

of those two figures. For the first figure, all models performed in a consistence way between ~90% to ~98% while it's dipping slowly from ~80% to ~50% in the next figure.

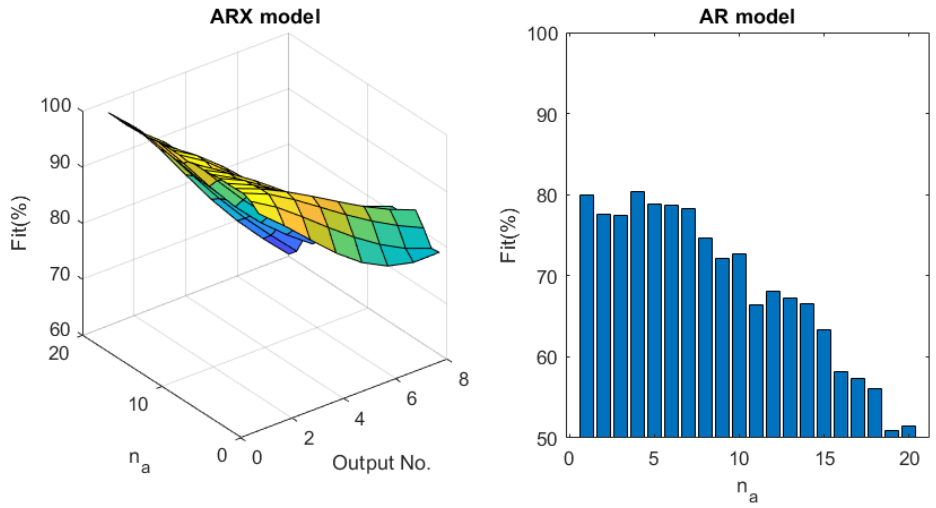

Figure 17. Performance of different ARX (left sub-figure) and AR (right sub-figure) models with the same training and validation data. In both cases, the parameter n_a is varying from 1 to 20.

Our training data has no information regarding the running segment's HRV, and the models failed to predict the output. A higher model order introduces more complexity in the model and results in a decreasing performance. At least we can say that we should avoid higher model orders for the AR model for our purpose, and the same goes true for the ARX model, though we need to define the correct sample shift for the input signal. The HRV signal is sampled at 8 Hz, and we should try to select a model in which the current output depends on at least 1 second of previous output, which refers to 8 sample points. For example, in the case of the ARX model, the model $ARX(12,1,0)$ means the model output depends on the current input and 12 previous outputs. We are assuming that using 12 previous output value will be enough to predict the current output. From now on, we will use the models $ARX(12,1,0)$ and $AR(12)$ for all the remaining experiments. For the sample shift in the input of the ARX model, if the sample shift is too low, we are feeding the current output to the input, and that doesn't make any sense. Again, if the sample shift is too high, we are feeding a non-relevant input, which might have no correlation to the output. We are assuming that 4 samples shift would be alright to serve our purpose. The reason for choosing the 4 samples shift is discussed later.

4.2. Effect of Different Training and Validation data

In the first running segment, the running speed is not fast enough to create a significant effect of running cadence on the heart rate variability. So far, we haven't used the HRV signal from other running segments yet. We were interested to see how the selected models will perform if we use the other running segment's HRV signal as validation data. Figure 18 shows the training and validation data from different running segments

for the model, and figure 19 demonstrates the performance of the model with different validation data.

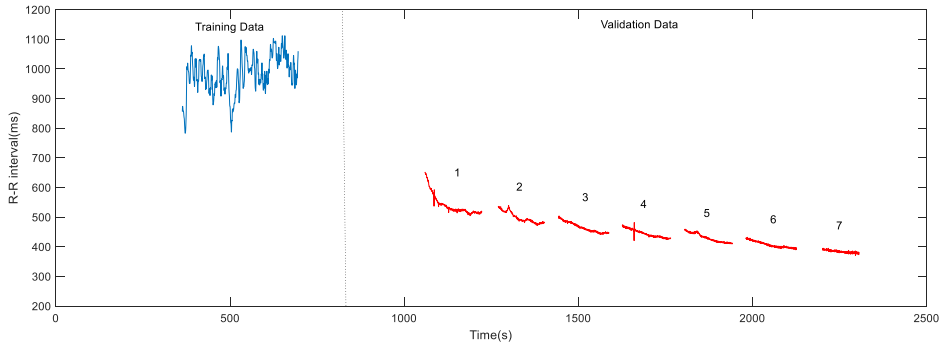

Figure 18. Training and validation data from participant ID 1 to observe the performance of the models $ARX(12,1,0)$ and $AR(12)$. The training data represents the HRV signal from the resting segment before running, and the validation data corresponds to the HRV signal from different running segments. The numeric values $(1 \dots 7)$ inside of the validation data stand for the running segment no.

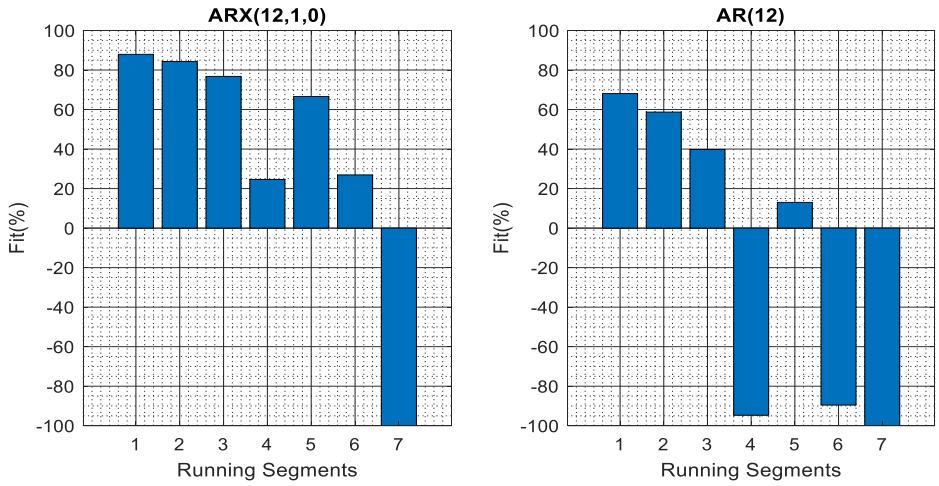

Figure 19. Bar representation of the performance of the models ARX(12,1,0) and AR(12)with different validation data from previous figure. Values beyond -100% is ignored.

By observing figure 19, we can say that there's an inverse relationship between the running speed and model performance if the model is trained with the HRV signal form normal resting condition. For the $ARX(12,1,0)$ model, the fit percentage is within the limit of $\sim(75\% - 85\%)$ when the first three running segments are used as the validation data. Then there's a sudden fall in fit percentage as it goes to $\sim(20\% - 25\%)$ for fourth and sixth running segments are used. When we are using the fifth running segment as the validation data, it doesn't strictly follow the inverse relation and shows an exceptional fit percentage of over $\sim 60\%$. For the last running segment, the worst result is expected, and the fit percentage goes beyond -100%. The fit percentages are truncated at -100% since we are not interested in negative fit percentages. For the $AR(12)$ model, the pattern of fit percentages is similar to the $ARX(12,1,0)$ model, though the actual values are different. The highest fit percentage

is approximately 70%, and the model shows a negative fit percentage for three running segments.

It's expected that model performance will decrease with the introduction of high running cadence in the validation data. This model is trained with the HRV signal from resting segments and from figure 18, we can see that the training data and validation data differ significantly. To achieve a better model performance and to build a more precise model, we should include some portion of the HRV signal from running condition in the training data.

We continued with the previous experiment with a modification in the training data. We were interested to see what will happen if we introduce the HRV signal from the running segment as training data. So, we used the first running segment's HRV as the training data. Figure 20 and figure 21 show the training and different validation data for the model and the model performance with the validation data, respectively.

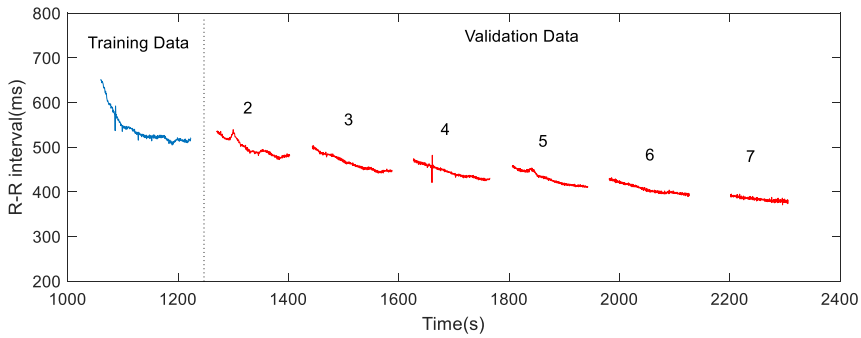

Figure 20. Training and validation data from participant ID 1 to observe the performance of the models $ARX(12,1,0)$ and $AR(12)$. The training data represents the HRV signal from the first running segment, and the validation data corresponds to the HRV signal from different running segments. The numeric values $(2 \dots 7)$ inside of the validation data stand for the running segment no.

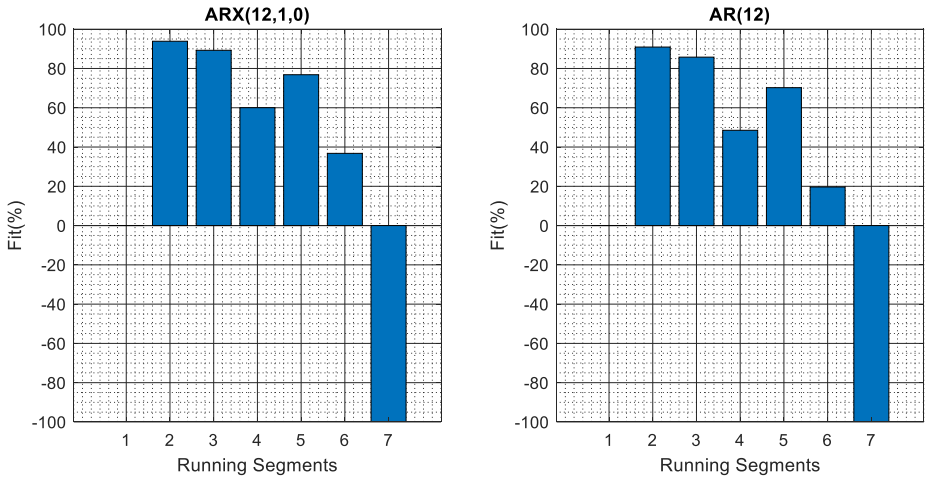

Figure 21. Bar representation of the performance of the models $ARX(12,1,0)$ and $AR(12)$with different validation data from the previous figure. Values beyond -100% are ignored.

The model performances in figure 21 are not satisfying, either, though we can see some significant improvement in some cases. Just replacing the training data by the running segments' HRV signal doesn't provide any dramatic positive impact. Still, both models have a negative fit percentage for the last running segment. We can assume that it's not appropriate to build the model by using the running segment's HRV signal only. We need to find a proper combination of the resting segment's HRV signal and running segment's HRV signal to train the model. We also need to define some different approaches to achieve better performance with that validation data. The selection of proper training data is discussed in the next section.

4.3. Training and Validation Data Selection

We have already assumed a proper model order, and now we are focusing on selecting the appropriate training data for our model. The dynamics of HRV signal change rapidly during exercise, and the selection of the training data must be done precisely. In an incremental running test, each running segment has a different impact on the HRV signal. During physical exercise, the sympathetic nervous system controls the regulation of heart rate. The increasing running speed results in an increased heart rate with low heart rate variability. The effect of running cadence on the HRV signal is also enhanced with the increment of running speed. So, it's quite unrealistic to assume that we can predict the HRV signal of high-speed running condition by using a model which is only trained by the normal condition's or any single running segment's HRV signal. Before going any further with the training data selection, it's worth to mention the change of mean signal level in different running intervals HRV signal. Table 3 states the change of mean HRV signal in every running segment.

Table 3. Mean HRV in the different running segments (ID 1).

Segment No	Speed (*km/h*)	Mean HRV (*ms*)
1	8	540.3
2	9	499.1
3	10	465.3
4	11	444,3
5	12	429.9
6	13	406.9
7	14	383.6

From table 3, we can see that the mean HRV from the last running segment is almost 30% less than the first running segment's mean HRV. This gradual change in every running segment's HRV forces us to rethink about how we will proceed further. As a result of that, we came up with a new hypothesis that we will use separate training data for separate running segments. This means each running segment has its' own individual model. We will proceed with this idea now and redefine it later.

In the earlier sub-chapter, we were discussing about the sample shift in the input of the ARX model, and we mentioned that 4 samples shift would be alright for our case. From table 3, the calculated mean HR of each segment are 111.05, 120.22, 128.95

$135.03, 139.57, 147.45, 156.41\ bpm,$ respectively. In case of intense exercise, the actual HR can easily go beyond $200\ bpm$. The HR of $120\ bpm$ is equivalent to $2\ bps,$ and due to our $8\ Hz$ sampling rate, one HR corresponds to 4 sample points in the HRV signal. After every 4 samples, we have a new HRV point. In the running segments, the mean HR is always over $120\ bpm$ or close to $120\ bpm$. By using 4 samples shift in the input, we are trying to make sure that we are feeding sample points close to the next HRV sample as input. We are trying to avoid feeding sample point close to current HRV as input since if the input and output is the same signal, there's no meaning of predicting the future value.

In the incremental running procedure, each running segment is followed by a resting segment. From now on, we are mentioning this segment as a resting interval between running segments to create a separation in naming from the initial resting segment. So far, we haven't used the resting interval's HRV signal anywhere. These resting intervals are not like the initial resting segment since there's a recovery from the training stress is going on those intervals. After each stage of running, the recovery is different, and that short resting interval is not enough for providing the full recovery. We can assume that a certain amount of information from running segments passes through the resting interval. If we use the resting interval's HRV signal as training data for the model, our model would be a perfect fit for the previous running segment. In our separate model structure, we trained each model with the resting interval's HRV signal and then validated the model with the previous running segment. The training and validation data for the separate model structure are given in figure 22, and the performance of each model is shown in figure 23.

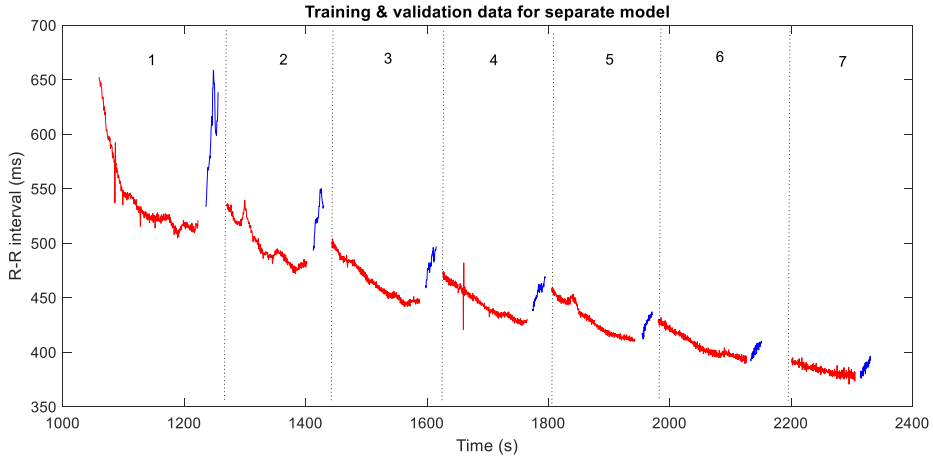

Figure 22. Separate training and validation data from participant ID 1 to observe the performance of the models $ARX(12,1,0)$ and $AR(12)$ for each segment. The colors, blue and red differentiate the training and validation data segment for each model. The numeric values $(1 \dots 7)$ stand for the model number.

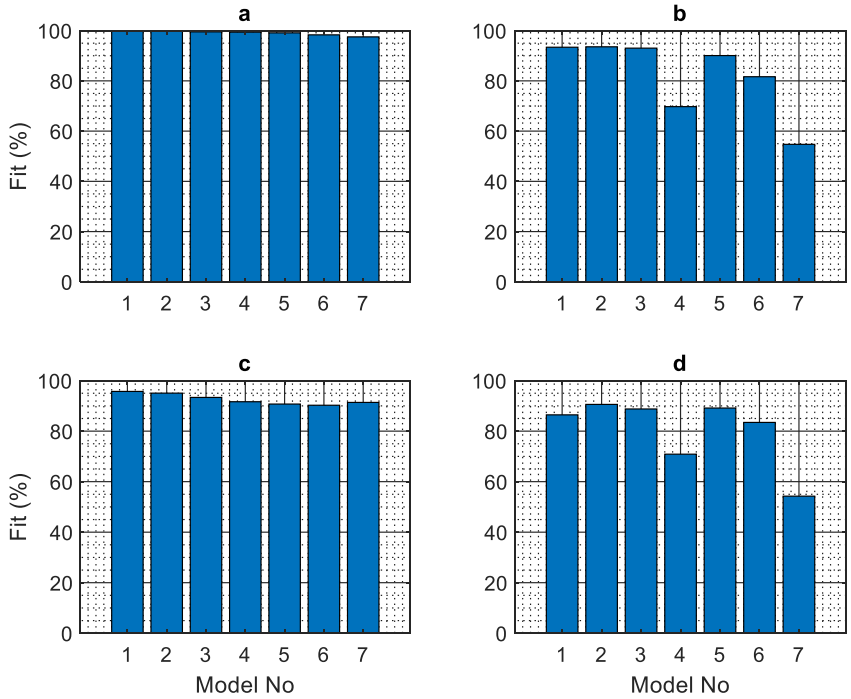

Figure 23. Bar representation of the performance of different models with different training and validation data. Sub-figures (a) and (b) show the fit percentage for $ARX(12,1,0)$ with the training and validation data, respectively. The last two sub-figures represent the same fit percentage for $AR(12)$.

The impact of separate models is clearly shown in figure 23. Separate models with the separate training data from figure 22 result a significant improvement in fit percentage with validation data for the later running segments (where the running speed is high). Since the models are unique, model performance with training data is also included in the figure to make sure that the models are properly trained. Another important observation is the similarity of the fit percentage for both $ARX(12,1,0)$ and $AR(12)$ models. Though the fit percentage with validation data is almost similar for both approaches, the fit percentage with validation data is a little different where the ARX approach provides a higher fit percentage with training data.

One issue with these separate models would be the lack of enough training data as the training data for each model is approximately $20\ seconds$ long, and the validation data has a much higher duration than the training data. Though we are assuming that the $20\ seconds$ training data would be enough for the model training, sometimes there's a possibility of wrong calculation of model parameters. The short training data can provide a higher fit percentage when the model is validated with the training data itself.

As designing separate models for separate running segments can be problematic, combining all the distinct models to a single model would be an excellent solution for that. For the next experiment, we tried to combine all the separate training data from figure 22 into a single training data and then used that training data to train a unique model. By combining all the training data, we assumed that the previous problem with the short training data would be gone, and our model would be more robust for all running segments. The performance of this single model is shown in figure 24.

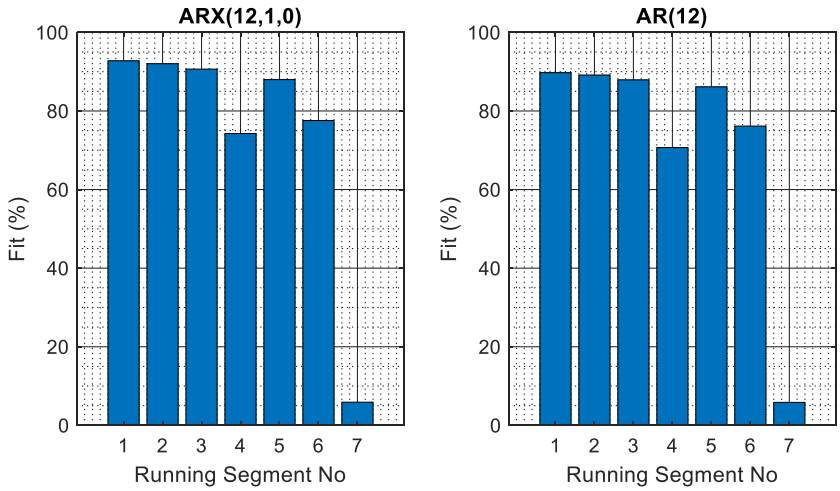

Figure 24. Bar representation of the performance of the models $ARX(12,1,0)$ and $AR(12)$ with different validation data from figure 22. Both models are trained with a combination of all training data from figure 22.

From figure 24, we can see that again we are facing the problem the with last running segment, i.e., the running segment with very high speed. Though our combined training data includes the resting interval after the last running segment, it includes all the other resting segments too, and that might increase the model complexity. Thus, the model fails to provide a better performance when the last running segment is used as the validation data. It's quite straight forward that we can't just combine all the training data. We need to find another approach to combine all the separate models from figure 23, which is done in the next experiment.

4.4. Chain Model Structure

For the next experiment, we have proposed an idea to combine the models rather than combining the training data. The main purpose is to create a chain of models where each model will be connected to each other. The idea of refining an existing model is used to create a chain model structure. First, we trained the initial model by using the first resting interval as the training data. Then this model was refined by using the following resting intervals one by one. Finally, the refining procedures ended when there are no resting segments are available. Thus, the final model had the information from all the resting segments, and the unnecessary complexities should be avoided as all the resting segments were not included at once. For validating the final model, we used the same validation data from figure 22. Figure 25 and figure 26 show the flowchart and the performance of the chain model, respectively.

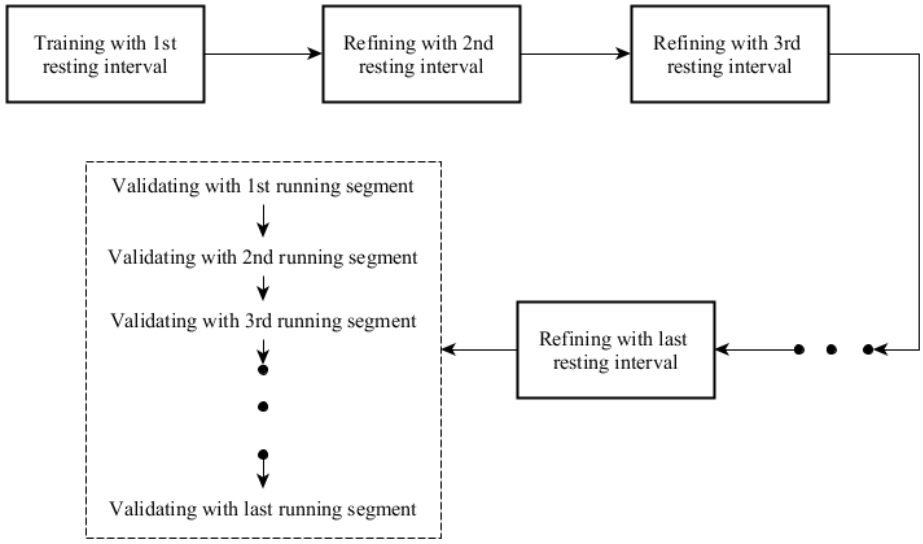

Figure 25. Flowchart of the proposed chain model system.

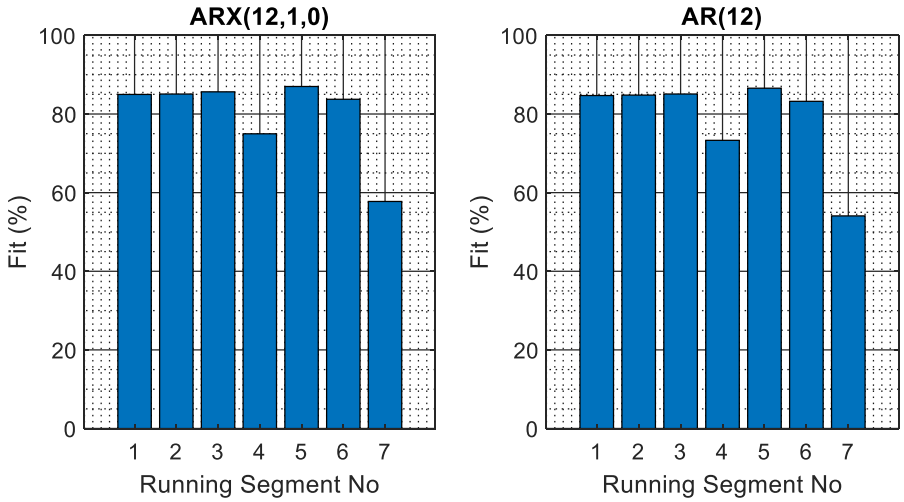

Figure 26. Bar representation of the performance of the chain models $ARX(12,1,0)$ and $AR(12)$ with different validation data from figure 22.

Though the model performances are a little bit lower for the earlier running segments, the overall model performance is relatively good. During comparing the actual output and the estimated output of the model, we found that this fit percentage values can be a little bit misleading. The problem lies in the refinement of the model. When we were refining the model with new training data, the model was trained again with the new training data. As a result, the final model of our chain was always trained with the last running interval as training data. Theoretically, the last resting interval should contain the information from all previous running intervals, and that's the reason for higher fit percentages, even for the first running segments. If we take a closer look at the original and estimated output signal from the chain model, the whole thing will be clearly visible. Figure 27 shows a couple of segments of the original and estimated output signal from both chain models, $ARX(12,1,0)$ and $AR(12)$.

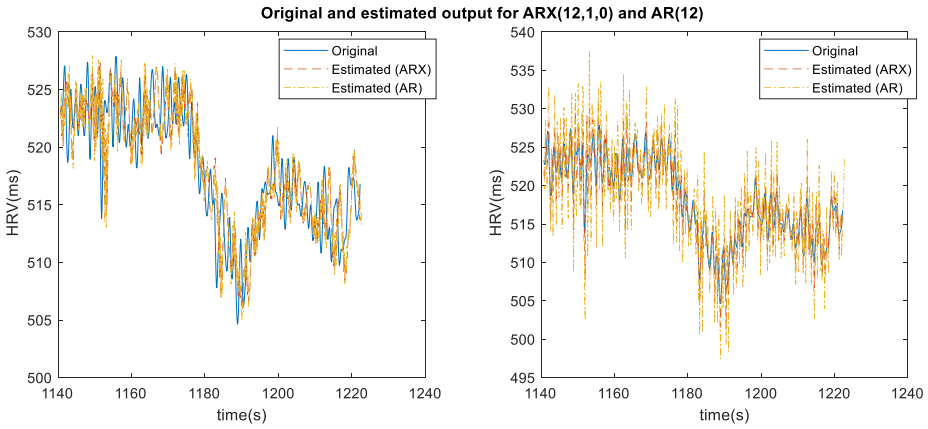

Figure 27. Original and estimated output from the chain model $ARX(12,1,0)$ and $AR(12)$, and the first running segment is used as the validation data. Only a portion of the validation data is shown in the figure. In the left sub-figure, the chain models are refined with the $7th$ resting interval while in the right sub-figure, the models are refined with the $1st$ resting interval.

From the left sub-figure of the above figure, it's visible that the estimated output is always shifted away from the original output when the last resting interval is used to refine the model and the first running segment as the validation data. This problem is solved if we use the first resting interval to refine the chain model. So, we can't continue with the idea of creating a single model for all running segments. We can create a single model, but before validating the model with running interval, we must refine the model with the appropriate resting interval. We need to add this slight adjustment in our chain model approach, and after this, the model performance is similar to figure 23. The updated flowchart of the chain model system is shown in figure 28. The original and estimated output for every running segment by using the chain models $ARX(12,1,0)$ and $AR(12)$ are shown in figure 29.

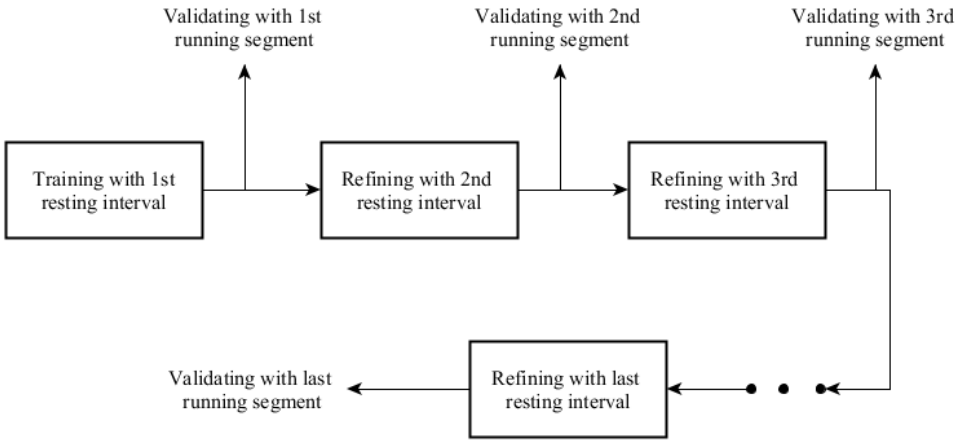

Figure 28. Flowchart of the adjusted chain model system.

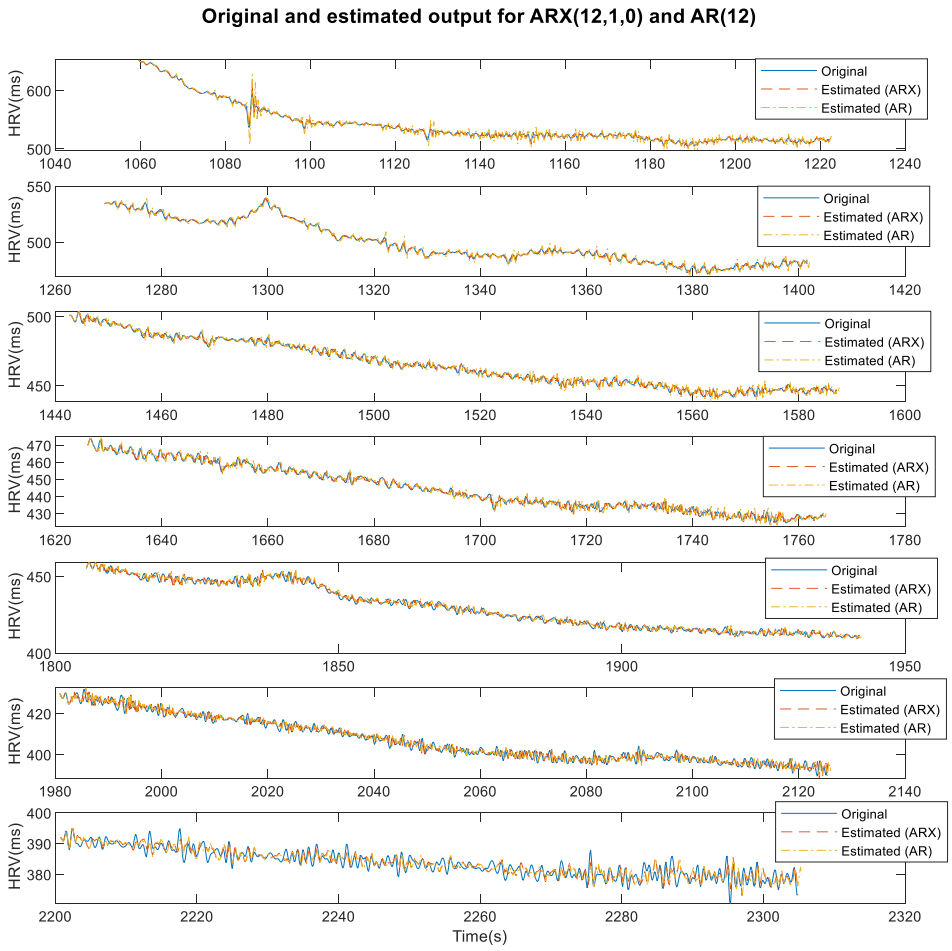

Figure 29. Original and estimated output from the chain model $ARX(12,1,0)$ and $AR(12)$. In the topmost sub-figure, the first running segment is used as the validation data, and the following running segments are used in the sub-figures, respectively.

4.5. Model Residuals

We are interested to see the error signal produced by the model since we assumed that the error signal might represent some useful information from the acceleration. The error signal is the difference between the actual HRV signal and the predicted HRV signal from the chain model. The error signal and autocorrelation of the error signal for $ARX(12,1,0)$ and $AR(12)$ are shown in figure 30 and figure 31, respectively.

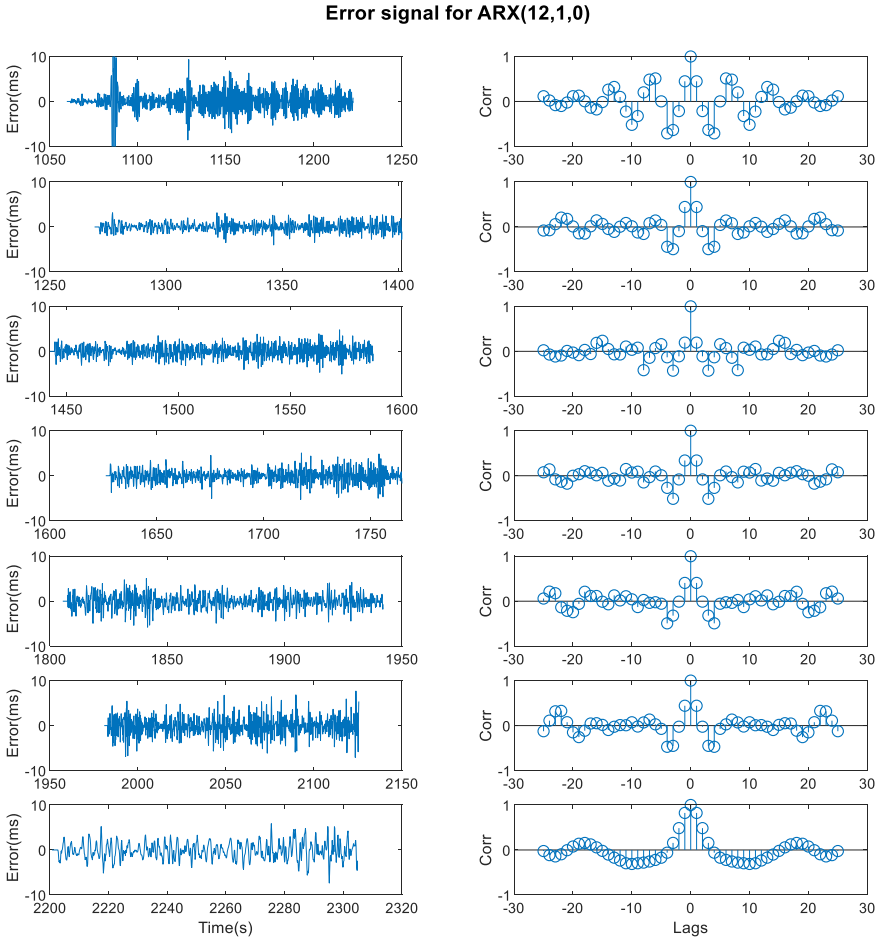

Figure 30. Error signal and the autocorrelation of the error signal for the chain model $ARX(12,1,0)$. Each row of the figure stands for each running segment, and it starts from the top to bottom in chronological order $(1,2,3,\ldots 7)$.

From figure 30 and figure 31, the autocorrelation of the error signal is similar for both $ARX(12,1,0)$ and $AR(12)$. Initially, the autocorrelation is random, and with the increment of running speed, the autocorrelation has become more periodic. This might refer to the effect of the acceleration on the HRV signal as we were expecting that this error signal should relate the accelerator signal.

The HRV signal varies from person to person, and a lot of factors like age, sex, exercise habit, etc. can affect it. With this type of differences between individuals, it's not possible to create a universal model which will be applicable for all participants. So far, the models were designed and validated by using the data from one participant. The modeling approach will be the same for all the participants, and we can create a distinct model for each participant by changing the training and validation data of the model. The general statistics of the model performance for the whole dataset are shown in table 4, and the actual model performance for every participant is included in the appendix.

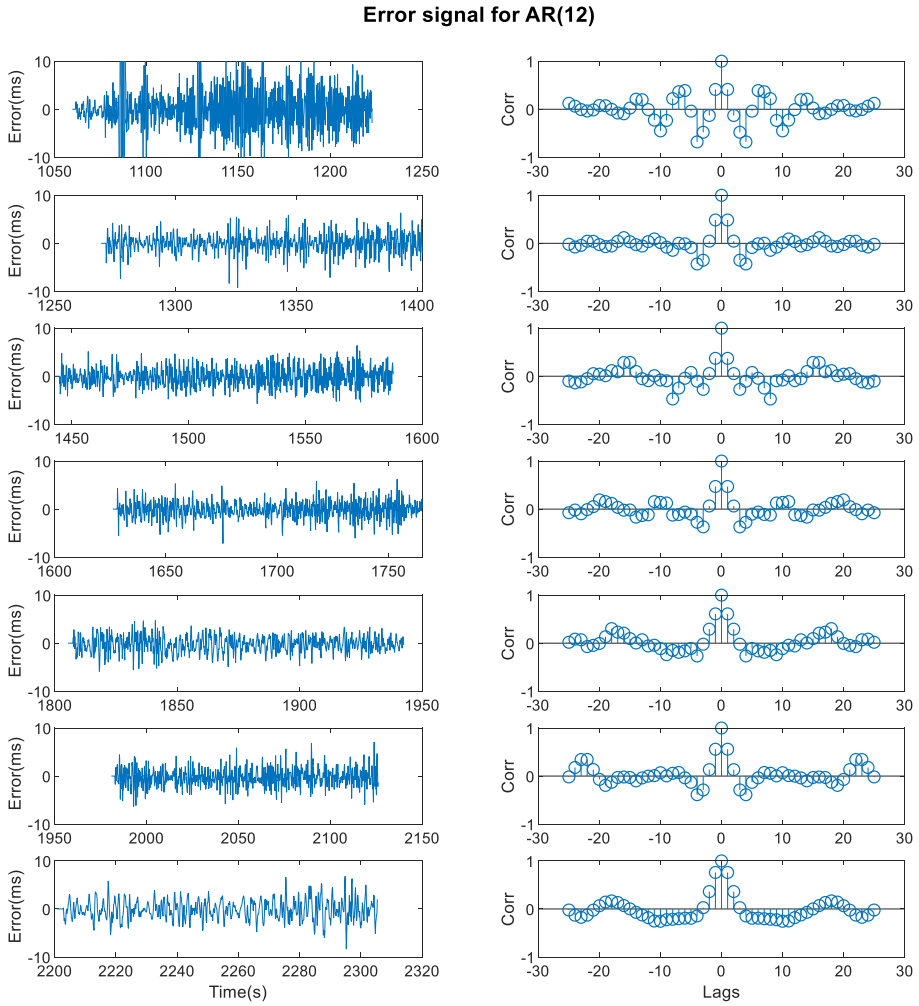

Figure 31. Error signal and the autocorrelation of the error signal for the chain model $AR(12)$. Each row of the figure stands for each running segment, and it starts from the top to bottom in chronological order (1,2,3, ... 7).

The number of participants varies for the later running segments since the participants continued running until exhaustion. Sometimes, the model performance decreases due to the poor signal quality, and that's why the standard deviation of the model performance increased. The model performance is usually worst for the last running segment.

After this calculation, the modeling process is finished, and in the next chapter, we will discuss the importance of the calculated error signal and the overall model performance with the whole dataset.

Table 4. General statistics of the model performances.

Running Segment No.	No. of Participants	Fit (%) $AR(12)$		Fit (%) $ARX(12,1,0)$	
		Mean	STD	Mean	STD
1	9	71.72	17.69	85.64	8.93
2	10	84.22	4.99	90.06	4.33
3	10	82.72	4.01	86.05	5.93
4	10	78.64	10.43	82.10	8.19
5	9	72.01	7.12	81.79	6.43
6	9	69.37	10.44	79.19	8.48
7	9	55.58	23.81	65.21	22.15
8	6	46.75	27.92	56.69	26.81
9	5	58.89	18.84	72.50	12.93
10	4	45.13	21.07	65.61	14.32

5. DISCUSSION

Our initial hypothesis was straight forward as we assumed that we could predict the current HRV signal by using the previous HRV signal. In the last chapter, we went through the system identification procedure to design such model or models which can provide the desired result. At the very beginning, we started with time series AR model for our purpose. The AR model has no input channel, and the HRV signal at a given point depends on the n previous samples where n is the model order. After observing various model order, we decided to use the model $AR(12)$. Besides the AR modeling, we were thinking about how to improvise our modeling approach.

Then we came up with the idea of providing an input signal which contains the value of the future HRV signal and thus creates a new modeling approach by using the ARX model. The input signal is generated by time shifting the original signal. For the ARX model, we chose the same model order as AR, and it's $ARX(12,1,0)$. This ARX model is using only one input sample at a time. We created several signals from the original signal by shifting the time from $0.125s$ to $1s$ and after many experiments, decided to use the $0.5s$ time shift in the input signal. So, if we are predicting a sample point of the HRV signal at a given time n, the $ARX(12,1,0)$ model is predicting it by using the sample point at $(n + 0.5)s$ in input and 12 previous output points.

The system identification procedure is based on trial and error, and we tried to include many different things; even all the kinds of stuff are not documented in this thesis. For example, we tried to improve the model performance by using model regularization. Since the results were not satisfactory, model regularization is excluded from our procedure. Similarly, we manipulated the training and validation data for the model several times, and only the useful ones are documented in the previous chapter.

In the extension of our previous hypothesis, we assumed that we could find some useful correlation between the model residual and the accelerometer signal. We also assumed that if we can feed the acceleration signal as the input of our model to create a multiple input single output (MISO) model. From this idea, we started to work with the ARX model structure rather than the conventional AR model structure. When we were trying to add another input channel for our $ARX(12,1,0)$ model, we recognized the huge difference in sampling rate between the acceleration signal and the HRV signal. In the used database, the acceleration signal is sampled at $1000\ Hz$ while we constructed the HRV signal at $8\ Hz$. This means the acceleration signal has 125 times higher sampling rate than the HRV signal. One quick solution would be down-sampling the acceleration signal to $8\ Hz$, which will create another problem since a lot of useful information will be lost due to this massive down-sampling.

Even if we accept the loss of information and down-sample the acceleration signal, the problem will not be fixed. Our current ARX model structure, $ARX(12,1,0)$ is not using any previous input values to calculate the output; rather, it's using the current input value only. The type of model structure won't be suitable for the MISO model since the input-output dependency can be different for different inputs. As a result, the redefinition of the model structure is required, and we need to go through the whole system identification procedure from the beginning.

Our initial goal was to design an HRV predictor model by using the dataset based on incremental running. From our designed model, we can see the gradual decrement of model performance with the running intensity. Signal quality is another crucial factor as sometimes the recorded ECG can be very bad due to the body movements. Especially in our used dataset, ID 6 has very bad signal quality, and for a couple of running segments, we can't use the modeling process. Also, in some cases, unusual artifacts are presented in the HRV signal that can affect model performance. We haven't set any acceptable limit for model performance though one large artifact can alter the model performance. For example, in the case of ID 6, a huge artifact was present in the HRV signal of one running segment, and that reduces the model performance to around $\sim 20\%$, where without the artifact, it's about $\sim 60\%$. To avoid confusion, we excluded these running segments from the calculation.

Our dataset included data from 10 participants. The first participant completed 7 running segments, and the signal quality was acceptable for all of those. We observed the gradual decreasing of the model performances as the fit percentage was $\sim 90\%$ for the first running segment and $\sim 55\%$ for the last running segment. The second participant followed the pattern of the first participant, and the fit percentage was further decreased to $\sim 30\%$ for the last running segment. The third and fourth participants completed 10 running segments. The fit percentage followed the gradual decreasing pattern though the fit percentage for the fourth participant was a little bit higher than the third participant. We observed a significant difference between the model performance of AR and ARX models for the fourth participant. The fifth participant completed 7 running segments. As we stated earlier, the signal quality for the sixth participant was poor; we excluded 6 of the completed 9 running segments of that participant. The seventh participant completed 8 running segments. After seeing the gradual decrement in model performances for the first 7 running segments, surprisingly, the last running segments showed a very good fit percentage of $\sim 85\%$. The eighth participant finished 10 running segments, and we removed one running segment for having a bad data quality. The ninth participant also completed the 10 running segments, and the fit percentages remained almost constant for all the running segments. We didn't observe the decreasing pattern for that participant. Finally, the last participant finished 9 running segments and showed the decreasing pattern in the model performances.

We haven't found any proper way to correlate the residual signal from the predictor model with the accelerometer signal. We are finishing with the assumption that by using the appropriate methods, we could find some kind of relationship between the residual and acceleration signal. This modeling process is still in the early phase, and there's a lot of scope for improvements. Though we couldn't fulfill all the hypotheses stated at the very beginning of the thesis, we went to a certain position from where this study can be continued later with further modifications and new ideas. We were also trying to create a non-linear model to include the non-linearities of HRV, but due to lack of experience and enough time, we failed to develop the desired non-linear model. We also studied different non-linear control techniques of the heart rate regulation, which would be useful for future work.

6. CONCLUSION

HRV analysis has become a popular research topic in the last few decades as it is a useful non-invasive technique for ANS assessment. Another exciting and promising research field in HRV analysis is the analysis of HRV during physical exercise. HRV dynamics change a lot during physical exercise, which makes the HRV analysis much more challenging. A lot of recent studies has shown the effect of different exercise like treadmill running, cycling on HRV.

During this study, we worked on a dataset based on incremental running on a treadmill. We tried to create a linear polynomial model which can predict the next HRV signal value based on the previous value. We introduced the AR and ARX models and defined two modeling approaches for our purpose. In the ARX approach, we included an HRV signal value from the near future to restrict our model to predict the HRV signal more accurately. Later, we calculated the model residuals from the assumption that the residual signal might correspond to the acceleration signal.

We created individual models for individual users since the HRV signal has its own dependency on age, sex, and other physiological matters. We validated the model by using HRV from different running segments and observed a gradual decrement in the model performance. We found a significant reduction in the model performance with higher running speed. For most of the participants, the model performance is worst when it's validating with the last running segments' HRV signal. The running speed is maximum during that running segment, and the effect of running cadence would be maximum too. We are assuming that the decrement in model performance is somehow related to the increasing impact of running cadence on the HRV signal.

Our initial idea was to introduce the acceleration signal in our system as a separate input. But due to the limitation in the pre-processing of the acceleration signal, we couldn't concatenate the acceleration signal to our model. The number of participants in our dataset is relatively small, and it would be better to validate this study again over a vast dataset. Our modeling approach is still in the initial phase, and it requires further modification and improvement.

7. REFERENCES

[1] Alikhani, I., Noponen, K., & Seppänen, T. (2017, July). Contribution of body movements on the heart rate variability during high intensity running. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3993-3996). IEEE.
DOI: http://dx.doi.org/10.1109/EMBC.2017.8037731

[2] Blain, G., Meste, O., Blain, A., & Bermon, S. (2009). Time-frequency analysis of heart rate variability reveals cardiolocomotor coupling during dynamic cycling exercise in humans. *American Journal of Physiology-Heart and Circulatory Physiology*, *296*(5), H1651-H1659.
DOI: http://dx.doi.org/10.1152/ajpheart.00881.2008

[3] Hernando, D., Hernando, A., Casajús, J. A., Laguna, P., Garatachea, N., & Bailón, R. (2018). Methodological framework for heart rate variability analysis during exercise: application to running and cycling stress testing. *Medical & biological engineering & computing*, *56*(5), 781-794.
DOI: http://dx.doi.org/10.1007/s11517-017-1724-9

[4] Manis, G., Nikolopoulos, S., Alexandridi, A., & Davos, C. (2007). Assessment of the classification capability of prediction and approximation methods for HRV analysis. *Computers in biology and medicine*, *37*(5), 642-654.
DOI: http://dx.doi.org/10.1016/j.compbiomed.2006.06.008

[5] Tarvainen, M. P., Georgiadis, S., & Karjalainen, P. A. (2006, January). Time-varying analysis of heart rate variability with kalman smoother algorithm. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (pp. 2718-2721). IEEE.
DOI: http://dx.doi.org/10.1109/IEMBS.2005.1617032

[6] Su, S. W., Huang, S., Wang, L., Celler, B. G., Savkin, A. V., Guo, Y., & Cheng, T. (2007, August). Nonparametric Hammerstein model based model predictive control for heart rate regulation. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 2984-2987). IEEE.
DOI: http://dx.doi.org/10.1109/IEMBS.2007.4352956

[7] Scalzi, S., Tomei, P., & Verrelli, C. M. (2011). Nonlinear control techniques for the heart rate regulation in treadmill exercises. *IEEE Transactions on Biomedical Engineering*, *59*(3), 599-603.
DOI: http://dx.doi.org/10.1109/TBME.2011.2179300

[8] Fox, S. I. (2006). *Human Physiology 8th Editon* (pp. 501-502). McGraw-Hill press, New York, USA.

[9] Jumilla, D. H. (2017). Non-Invasive Assessment of Autonomic Nervous System through Cardiovascular Signals Variability Analysis in Non-Stationary Environments. Ph. D. Thesis. University of Zaragoza. Department of Biomedical Engineering.

[10] McCorry, L. K. (2007). Physiology of the autonomic nervous system. *American journal of pharmaceutical education*, *71*(4), 78.
DOI: http://dx.doi.org/10.5688/aj710478

[11] Hall, J. E. (2015). *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences.

[12] Sztajzel, J. (2004). Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss medical weekly*, *134*(35-36), 514-522.
DOI: http://dx.doi.org/2004/35/smw-10321

[13] James, T. N. (2003). Structure and function of the sinus node, AV node and his bundle of the human heart: part II--function. *Progress in cardiovascular diseases*, *45*(4), 327.
DOI: http://dx.doi.org/10.1053/pcad.2003.1

[14] Rangayyan, R. M. (2015). *Biomedical signal analysis* (Vol. 33). John Wiley & Sons.

[15] Sörnmo, L., & Laguna, P. (2005). *Bioelectrical signal processing in cardiac and neurological applications* (Vol. 8). Academic Press.

[16] Electrocardiogram (ECG): working principle, normal ECG wave, application of ECG, https://www.onlinebiologynotes.com/electrocardiogram-ecg-working-principle-normal-ecg-wave-application-of-ecg/

[17] Malik, M. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European Society of Cardiology and the North American Society for Pacing and Electrophysiology. *Annals of Noninvasive Electrocardiology*, *1*(2), 151-181.

[18] Huikuri, H. V., Mäkikallio, T., Airaksinen, K. J., Mitrani, R., Castellanos, A., & Myerburg, R. J. (1999). Measurement of heart rate variability: a clinical tool or a research toy?. *Journal of the American College of Cardiology*, *34*(7), 1878-1883.
DOI: http://dx.doi.org/10.1016/s0735-1097(99)00468-4

[19] Dreifus, L. S., Agarwal, J. B., Botvinick, E. H., Ferdinand, K. C., Fisch, C., Fisher, J. D., ... & Prystowsky, E. N. (1993). Heart rate variability for risk stratification of

life-threatening arrhythmias. *Journal of the American College of Cardiology*, *22*(3), 948-950.
DOI: http://dx.doi.org/10.1016/0735-1097(93)90217-O

[20] Hohnloser, S. H., Klingenheben, T., Zabel, M., & Li, Y. G. (1997). Heart rate variability used as an arrhythmia risk stratifier after myocardial infarction. *Pacing and clinical electrophysiology*, *20*(10), 2594-2601.
DOI: http://dx.doi.org/10.1111/j.1540-8159.1997.tb06109.x

[21] Barron, H. V., & Viskin, S. (1998). Autonomic markers and prediction of cardiac death after myocardial infarction. *The Lancet*, *351*(9101), 461-462.
DOI: http://dx.doi.org/10.1016/S0140-6736(05)78676-1

[22] Wijbenga, J. A. M., Balk, A. H. M. M., Meij, S. H., Simoons, M. L., & Malik, M. (1998). Heart rate variability index in congestive heart failure: relation to clinical variables and prognosis. *European heart journal*, *19*(11), 1719-1724.
DOI: http://dx.doi.org/10.1053/euhj.1998.1148

[23] Rothschild, M., Rothschild, A., & Pfeifer, M. (1988). Temporary decrease in cardiac parasympathetic tone after acute myocardial infarction. *The American journal of cardiology*, *62*(9), 637-639.
DOI: http://dx.doi.org/10.1016/0002-9149(88)90670-4

[24] Melillo, P., Bracale, M., & Pecchia, L. (2011). Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination. *Biomedical engineering online*, *10*(1), 96.
DOI: http://dx.doi.org/10.1186/1475-925X-10-96

[25] Salahuddin, L., & Kim, D. (2006, November). Detection of acute stress by heart rate variability (HRV) using a prototype mobile ECG sensor. In *Proceedings of the International Conference on Hybrid Information Technology, Cheju Island, Korea* (pp. 9-11).

[26] Vandeput, S., Taelman, J., Spaepen, A., & Van Huffel, S. (2009, June). Heart rate variability as a tool to distinguish periods of physical and mental stress in a laboratory environment. In *Proceedings of the 6th international workshop on biosignal interpretation (BSI), New Haven, CT* (pp. 187-190).

[27] Sun, F. T., Kuo, C., Cheng, H. T., Buthpitiya, S., Collins, P., & Griss, M. (2010, October). Activity-aware mental stress detection using physiological sensors. In *International conference on Mobile computing, applications, and services* (pp. 282-301). Springer, Berlin, Heidelberg.
DOI: http://dx.doi.org/10.1007/978-3-642-29336-8_16

[28] Kujala, U., Pietilä, J., Myllymäki, T., Mutikainen, S., Föhr, T., Korhonen, I., & Helander, E. (2017). Physical activity: Absolute intensity vs. relative-to-fitness-level volumes. *Medicine and science in sports and exercise*, *49*.
DOI: http://dx.doi.org/10.1249/MSS.0000000000001134

[29] Hynynen, E., Konttinen, N., Kinnunen, U., Kyröläinen, H., & Rusko, H. (2011). The incidence of stress symptoms and heart rate variability during sleep and orthostatic test. *European journal of applied physiology*, *111*(5), 733-741. DOI: http://dx.doi.org/10.1007/s00421-010-1698-x

[30] Plews, D. J., Laursen, P. B., Stanley, J., Kilding, A. E., & Buchheit, M. (2013). Training adaptation and heart rate variability in elite endurance athletes: opening the door to effective monitoring. *Sports medicine*, *43*(9), 773-781. DOI: http://dx.doi.org/10.1007/s40279-013-0071-8

[31] Umetani, K., Singer, D. H., McCraty, R., & Atkinson, M. (1998). Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. *Journal of the American College of Cardiology*, *31*(3), 593-601. DOI: http://dx.doi.org/10.1016/S0735-1097(97)00554-8

[32] Yamasaki, Y., Kodama, M., Matsuhisa, M., Kishimoto, M., Ozaki, H., Tani, A., ... & Kamada, T. (1996). Diurnal heart rate variability in healthy subjects: effects of aging and sex difference. *American Journal of Physiology-Heart and Circulatory Physiology*, *271*(1), H303-H310. DOI: http://dx.doi.org/10.1152/ajpheart.1996.271.1.H303

[33] Ljung, L. (2001). System identification. *Wiley Encyclopedia of Electrical and Electronics Engineering*.

[34] Keesman, K. J. (2011). *System identification: an introduction*. Springer Science & Business Media.

[35] Ljung, L. (1995). *System identification toolbox: User's guide*. Natick, MA: MathWorks Incorporated.

[36] Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng*, *32*(3), 230-236. DOI: http://dx.doi.org/10.1109/TBME.1985.325532

[37] Micheal Walden, Heartbeat and cardiac conduction system. https://www.teachpe.com/anatomy-physiology/heart-rate-cardiac-volumes

[38] Introduction to cardiac excitation & contraction. http://tmedweb.tulane.edu/pharmwiki/doku.php/introduction_to_cardiac_physiology_electrophysiology

8. APPENDIX

Table 5. Model performance for the whole dataset

ID \ RS		1	2	3	4	5	6	7	8	9	10
1	ar	86.44	90.39	88.82	87.11	89	83.42	54.08			
	arx	93.23	94.7	92.48	89.75	89.65	82.1	57.76			
2	ar	83.75	86.68	84.19	77.12	65.48	59.42	31.25			
	arx	91.53	88.7	84.05	80.38	76.88	72.94	28.13			
3	ar	71.01	87.68	87	85.69	80.94	80.71	80.88	71.86	70.78	44.2
	arx	86.39	93.91	92.56	92.52	88.58	86.69	80.48	74.37	67.12	49.45
4	ar	85.87	85.68	86.15	87.47	84.26	86.17	82.9	84.26	76.41	57.29
	arx	93.06	92.7	92.59	89.82	86.04	86.08	83.61	88.02	85.76	78.42
5	ar	34.77	83.62	78.05	75.43	70.08	59.2	29.27			
	arx	64.25	90.06	78.46	75.44	68.85	60.71	29.48			
6	ar	--	86.37	83.3	50.09	--	--	--	--	--	
	arx	--	93.3	90.07	65.23	--	--	--	--	--	
7	ar	74.22	70.76	77.46	76.04	78.58	69.35	28.36	31.08	81.62	
	arx	85.75	80.42	77.68	74.48	78.62	72.43	56.05	30.69	89.34	
8	ar	74.91	82.27	76.4	79.84	80.25	81.79	78.09	26.94	--	40.6
	arx	86.85	84.49	78.72	80.25	80.06	82.22	81.91	29.55	--	53.32
9	ar	46.83	85.06	83.83	83.59	86.92	88.55	87.03	85.59	83.31	73.55
	arx	77.42	92.74	90.25	88.91	87.82	89.12	86.9	87.02	86.84	81.24
10	ar	87.67	83.73	82.03	84.03	79.59	79.12	76.91	19.51	32.23	
	arx	92.27	89.54	83.66	84.2	79.61	80.39	82.61	30.47	33.45	